

List of potentially matching sequences

Send selected sequences to Clustal W (multiple alignment)Submit QuerySelect up to...☐ Include query sequence

Db AC	Description	Score	E-value
<input type="checkbox"/> sp P19196	INVA_YEREN <u>Invasin</u> <u>[Yersinia enterocolitica]</u>	1639	0.0
<input type="checkbox"/> tr Q56889	Invasin [invA] [Yersinia enterocolitica]	1613	0.0
<input type="checkbox"/> sp P11922	INVA_YERPS Invasin [Yersinia pseudotuberculosis]	726	0.0
<input type="checkbox"/> tr Q56937	Invasin [inv] [Yersinia pestis]	721	0.0
<input type="checkbox"/> tr Q8ZA73	Putative invasin (Hypothetical) [YPO3944] [Yersinia pe...	428	e-118
<input type="checkbox"/> tr Q74QZ6	Putative invasin [YP3306] [Yersinia pestis]	428	e-118
<input type="checkbox"/> tr Q6KC27	Intimin lambda <u>[eae]</u> [Escherichia coli]	342	2e-92
<input type="checkbox"/> tr Q8KRK8	Intimin lambda [Escherichia coli]	340	6e-92
<input type="checkbox"/> tr Q8VL95	Theta <u>intimin</u> [eae] [Escherichia coli]	336	1e-90
<input type="checkbox"/> tr Q8VL00	Theta intimin [Escherichia coli]	335	2e-90
<input type="checkbox"/> sp O31000	EAE_ECO11 Intimin (Attaching and effacing protein) (Ea...	333	6e-90
<input type="checkbox"/> tr Q6EV99	Intimin sigma [eae] [Escherichia coli]	332	2e-89
<input type="checkbox"/> tr Q93FK0	Intimin <u>[Citrobacter rodentium]</u>	331	3e-89
<input type="checkbox"/> tr Q93UI3	Intimin type gamma [eae] [Escherichia coli]	331	4e-89
<input type="checkbox"/> tr Q842D8	Intimin [eae] [Escherichia coli]	330	5e-89
<input type="checkbox"/> tr P77067	Beta intimin (Eae) [eaeA] [Escherichia coli]	330	5e-89
<input type="checkbox"/> tr Q9RCH1	Intimin [eaeA] [Escherichia coli]	330	7e-89
<input type="checkbox"/> tr Q9R8R5	Intimin [eae] [Escherichia coli]	330	7e-89
<input type="checkbox"/> tr Q9S6R2	<u>Intimin</u> [eaeA] [Escherichia coli]	330	7e-89
<input type="checkbox"/> tr Q6KC44	Intimin mu [eae] [Escherichia coli]	330	7e-89
<input type="checkbox"/> tr Q47018	Intimin [eaeA] [Escherichia coli]	330	9e-89
<input type="checkbox"/> tr Q8X6G3	Putative adhesin [eaeH] [Escherichia coli O157:H7]	329	1e-88
<input type="checkbox"/> tr Q93K93	Intimin-kappa [eae-kappa] [Escherichia coli]	329	1e-88
<input type="checkbox"/> tr Q7AHB0	Putative invasin [ECs0336] [Escherichia coli O157:H7]	329	1e-88
<input type="checkbox"/> tr Q8FKK3	Putative adhesin [eaeH] [Escherichia coli O6]	329	1e-88
<input type="checkbox"/> tr Q93K94	Intimin-jota [eae-jota] [Escherichia coli]	328	2e-88
<input type="checkbox"/> tr Q9LBG6	Intimin [eae] [Escherichia coli]	328	2e-88
<input type="checkbox"/> tr Q6KC43	Intimin nu [eae] [Escherichia coli]	328	2e-88
<input type="checkbox"/> tr Q8KRL1	Intimin epsilon 2 [Escherichia coli]	327	4e-88
<input type="checkbox"/> tr Q9EXN4	Intimin zeta [eae] [Escherichia coli]	326	1e-87
<input type="checkbox"/> tr Q8KRK9	Intimin beta 2 [eae] [Escherichia coli]	326	1e-87
<input type="checkbox"/> sp Q07591	<u>EAE</u> CITFR Intimin (Attaching and effacing protein) (Ea...	326	1e-87
<input type="checkbox"/> tr Q6QZW1	Eae~ (Fragment) [Escherichia coli]	326	1e-87
<input type="checkbox"/> tr Q6KC41	Intimin pi [eae] [Escherichia coli]	326	1e-87
<input type="checkbox"/> tr Q6KC42	Intimin xi [eae] [Escherichia coli]	325	2e-87
<input type="checkbox"/> tr Q8VL93	Zeta intimin [Escherichia coli]	325	2e-87

<input type="checkbox"/>	tr Q846P2	Intimin-zeta [eae] [Escherichia coli]	325	3e-87
<input type="checkbox"/>	sp P43261	EAE_ECO57 Intimin (Attaching and effacing protein) (Ea...	324	4e-87
<input type="checkbox"/>	tr Q9RGP3	Intimin type epsilon [eae] [Escherichia coli]	324	4e-87
<input type="checkbox"/>	tr Q9R8B5	Gamma intimin [eae] [Escherichia coli]	324	4e-87
<input type="checkbox"/>	tr Q9R8B4	Gamma intimin [eae] [Escherichia coli]	324	4e-87
<input type="checkbox"/>	tr Q8RSS9	Intimin eta [eae] [Escherichia coli]	323	1e-86
<input type="checkbox"/>	tr Q84FQ2	Intimin epsilon [Escherichia coli]	323	1e-86
<input type="checkbox"/>	tr Q93UI4	Intimin type beta [eae] [Escherichia coli]	322	2e-86
<input type="checkbox"/>	tr Q8RNT8	Intimine type epsilon [eae] [Escherichia coli]	322	2e-86
<input type="checkbox"/>	tr Q8KRL0	Intimin alpha 2 [eae] [Escherichia coli]	321	4e-86
<input type="checkbox"/>	tr O06896	Intimin [eae] [Escherichia coli]	318	2e-85
<input type="checkbox"/>	tr O52620	Intimin [eaeA] [Escherichia coli]	318	3e-85
<input type="checkbox"/>	sp P19809	EAE_ECO27 Intimin (Attaching and effacing protein) (Ea...	317	6e-85
<input type="checkbox"/>	tr Q9F609	Intimin type epsilon (Fragment) [eaeA] [Escherichia coli]	316	1e-84
<input type="checkbox"/>	tr Q8KRL2	Intimin jota 2 [Escherichia coli]	316	1e-84
<input type="checkbox"/>	tr Q6KAV7	Intimin rho [eae] [Escherichia coli]	311	2e-83
<input type="checkbox"/>	tr Q7N599	Similarities with putative adhesin [plu2057] [Photorha...	310	9e-83
<input type="checkbox"/>	tr Q9EYM6	Intimin type epsilon (Fragment) [eaeA] [Escherichia coli]	310	9e-83
<input type="checkbox"/>	tr Q8XB95	Putative invasin [z5932] [Escherichia coli O157:H7]	304	4e-81
<input type="checkbox"/>	tr Q7A8L6	Putative invasin [ECs5290] [Escherichia coli O157:H7]	298	4e-79
<input type="checkbox"/>	tr Q8D083	Putative adhesin [y2605] [Yersinia pestis]	296	1e-78
<input type="checkbox"/>	sp P76347	YEEJ_ECOLI Hypothetical protein yeeJ [yeeJ] [Escherich...	293	1e-77
<input type="checkbox"/>	sp Q8X8V7	YEEJ_ECO57 Hypothetical protein yeeJ [z3135] [Escheric...	289	2e-76
<input type="checkbox"/>	tr Q7CQ17	SinH [sinH] [Salmonella typhimurium]	245	3e-63
<input type="checkbox"/>	tr Q9XCI9	SinH [sinH] [Salmonella typhimurium]	245	3e-63
<input type="checkbox"/>	tr Q8FF61	SinH homolog [c3031] [Escherichia coli O6]	235	3e-60
<input type="checkbox"/>	tr Q9L574	Invasin-like SivH [sivH] [Salmonella typhi]	233	9e-60
<input type="checkbox"/>	tr Q7X2C2	Aec1 precursor [aec1] [Escherichia coli]	233	1e-59
<input type="checkbox"/>	sp P36943	EAHE_ECOLI Attaching and effacing protein homolog prec...	209	1e-52
<input type="checkbox"/>	sp P39165	YCHO_ECOLI Hypothetical protein ychO [ychO] [Escherich...	199	2e-49
<input type="checkbox"/>	tr Q8XDD4	YchP protein [ychP] [Escherichia coli O157:H7]	198	3e-49
<input type="checkbox"/>	tr Q7AEU6	Putative factor [ECs1725] [Escherichia coli O157:H7]	198	3e-49
<input type="checkbox"/>	tr Q83RN8	Putative factor [ychP] [Shigella flexneri]	198	4e-49
<input type="checkbox"/>	tr Q7UCS2	Putative factor [ychP] [Shigella flexneri]	198	4e-49
<input type="checkbox"/>	tr Q8FHZ3	Hypothetical protein ychP [ychP] [Escherichia coli O6]	195	3e-48
<input type="checkbox"/>	tr Q93GR3	Putative adhesin [PSLT034] [Salmonella typhimurium]	190	8e-47
<input type="checkbox"/>	tr Q8ZP33	Putative invasin [ychP] [Salmonella typhimurium]	186	1e-45
<input type="checkbox"/>	tr Q8Z7G3	Putative invasin [STY1284] [Salmonella typhi]	186	1e-45
<input type="checkbox"/>	tr Q8ZP78	Homology to invasin C of Yersinia; intimin [STM1669] [...]	158	5e-37
<input type="checkbox"/>	tr Q83T27	Invasin-like protein [t1573] [Salmonella typhi]	157	1e-36
<input type="checkbox"/>	tr Q8Z795	Invasin-like protein [STY1395] [Salmonella typhi]	156	1e-36
<input type="checkbox"/>	tr Q7WR47	Putative adhesin [BB0110] [Bordetella bronchiseptica (...]	85	5e-15
<input type="checkbox"/>	tr Q7W286	Putative adhesin [BPP0104] [Bordetella parapertussis]	85	6e-15
<input type="checkbox"/>	tr Q7WAX9	Putative outer membrane ligand binding protein [bipA] ...	70	1e-10
<input type="checkbox"/>	tr Q9APE8	Putative outer membrane ligand binding protein [bipA] ...	69	4e-10

<input type="checkbox"/>	tr	Q7VZ27	Putative outer membrane ligand binding protein [bipA] ...	69	4e-10
<input type="checkbox"/>	tr	Q7VR49	Putative adhesin [eaeH] [Candidatus Blochmannia florid...	67	1e-09
<input type="checkbox"/>	tr	Q52708	Outer membrane protein A (Fragment) [ompA] [Rickettsia...	53	3e-05
<input type="checkbox"/>	tr	Q6CD35	Similarity [YALI0C04136g] [Yarrowia lipolytica CLIB99]	53	3e-05
<input type="checkbox"/>	tr	Q82XT8	Hemolysin-type calcium-binding region:RTX N-terminal d...	51	1e-04
<input type="checkbox"/>	sp	Q52657	OMPA_RICCN Outer membrane protein A precursor (190 kDa...	50	2e-04
<input type="checkbox"/>	tr	Q6CPZ4	Kluyveromyces lactis strain NRRL Y-1140 chromosome E o...	50	2e-04
<input type="checkbox"/>	sp	Q8TFG9	YL61_SCHPO Hypothetical serine/threonine-rich protein ...	47	0.002
<input type="checkbox"/>	tr	Q9C105	SPAPB1E7.04c protein [SPAPB1E7.04c] [Schizosaccharomyc...	47	0.002
<input type="checkbox"/>	tr	Q722C9	Cell wall surface anchor family protein [LMOF2365_0805...	46	0.003
<input type="checkbox"/>	tr	Q76I98	Bacterial immunogloblin [is2-2] [Leptospira interrogan...	46	0.003
<input type="checkbox"/>	tr	Q7U7M8	Hypothetical [SYNW0953] [Synechococcus sp. (strain WH8...	45	0.004
<input type="checkbox"/>	tr	Q8FGD9	Hypothetical protein c2439 [c2439] [Escherichia coli O6]	45	0.006
<input type="checkbox"/>	tr	Q6KB38	Intimin zeta (Fragment) [eae] [Escherichia coli]	45	0.006
<input type="checkbox"/>	tr	Q6AHX5	Hypothetical protein DKFZp686O19206 [DKFZp686O19206] [...	45	0.006
<input type="checkbox"/>	sp	P15921	OMPA_RICRI Outer membrane protein A precursor (190 kDa...	45	0.007
<input type="checkbox"/>	tr	Q8VVW0	Intimin-lambda (Fragment) [eae-lambda] [Escherichia coli]	44	0.010
<input type="checkbox"/>	tr	Q8VLL9	Intimin (Fragment) [eae] [Escherichia coli]	44	0.010
<input type="checkbox"/>	tr	Q6KBV3	Intimin zeta (Fragment) (Fragment) [eae] [Escherichia ...	44	0.010

sp P19196
INVA YEREN

835 AA
align

Query:	1	MYSFFNTLTVTKXXXXXXXXXXXXXXXXXXTYGFSQQHYFNSEALENPAEHNEAFNKIISTG	60
		MYSFFNTLTVTK	TYGFSQQHYFNSEALENPAEHNEAFNKIISTG
Sbjct:	1	MYSFFNTLTVTKIISRLILSIGLIFGIFTYGFSQQHYFNSEALENPAEHNEAFNKIISTG	60
Query:	61	TSLAVSGNASNITRSMVNDAANQEVKHWLNRFGTTQVNVNFDKKFSLKESSLDWLLPWYD	120
		TSLAVSGNASNITRSMVNDAANQEVKHWLNRFGTTQVNVNFDKKFSLKESSLDWLLPWYD	
Sbjct:	61	TSLAVSGNASNITRSMVNDAANQEVKHWLNRFGTTQVNVNFDKKFSLKESSLDWLLPWYD	120
Query:	121	SASYVFFSQLGIRNKDSRNTLNIGAGVRTFQQSWMYGFNTSYDNDMTGHNHRIGVGAEAW	180
		SASYVFFSQLGIRNKDSRNTLNIGAGVRTFQQSWMYGFNTSYDNDMTGHNHRIGVGAEAW	
Sbjct:	121	SASYVFFSQLGIRNKDSRNTLNIGAGVRTFQQSWMYGFNTSYDNDMTGHNHRIGVGAEAW	180
Query:	181	TDYLQLSANGYFRLNGWHQSRDFADYNERPASGGDIHVKAYLPALPQLGGKLYEQYRGE	240
		TDYLQLSANGYFRLNGWHQSRDFADYNERPASGGDIHVKAYLPALPQLGGKLYEQYRGE	
Sbjct:	181	TDYLQLSANGYFRLNGWHQSRDFADYNERPASGGDIHVKAYLPALPQLGGKLYEQYRGE	240
Query:	241	RVALFGKDNLQSNPYAVTTGLIYTPIPFITLQVDQRMGKSRQHEIQWNLQMDYRLGESFR	300
		RVALFGKDNLQSNPYAVTTGLIYTPIPFITLQVDQRMGKSRQHEIQWNLQMDYRLGESFR	
Sbjct:	241	RVALFGKDNLQSNPYAVTTGLIYTPIPFITLQVDQRMGKSRQHEIQWNLQMDYRLGESFR	300
Query:	301	SQFSPAVVAGTRLLAESRYNLVERNPNIVLEYQKQNTIKLAFSPAVLSGLPGQVYSVSAQ	360
		SQFSPAVVAGTRLLAESRYNLVERNPNIVLEYQKQNTIKLAFSPAVLSGLPGQVYSVSAQ	
Sbjct:	301	SQFSPAVVAGTRLLAESRYNLVERNPNIVLEYQKQNTIKLAFSPAVLSGLPGQVYSVSAQ	360
Query:	361	IQSQSALQRIWLWDAQWVAAGGKLIPVSATDYNVVLPPYKPMAPASRTVGKTGESEAAVN	420
		IQSQSALQRIWLWDAQWVAAGGKLIPVSATDYNVVLPPYKPMAPASRTVGKTGESEAAVN	
Sbjct:	361	IQSQSALQRIWLWDAQWVAAGGKLIPVSATDYNVVLPPYKPMAPASRTVGKTGESEAAVN	420
Query:	421	TYTLSATAIDNHGSSNPATLTVIVQQPQFVITSEVTDGALADGRTPITVKFTVTNIDS	480
		TYTLSATAIDNHGSSNPATLTVIVQQPQFVITSEVTDGALADGRTPITVKFTVTNIDS	
Sbjct:	421	TYTLSATAIDNHGSSNPATLTVIVQQPQFVITSEVTDGALADGRTPITVKFTVTNIDS	480
Query:	481	TPVAEQEGVITTSNGALPSKVTKKTDAGQVISIALTSFTVGVSVVTLDIQGQQATVDVRF	540
		TPVAEQEGVITTSNGALPSKVTKKTDAGQVISIALTSFTVGVSVVTLDIQGQQATVDVRF	
Sbjct:	481	TPVAEQEGVITTSNGALPSKVTKKTDAGQVISIALTSFTVGVSVVTLDIQGQQATVDVRF	540
Query:	541	AVLPPDVTNSSFNVSPSDIVADGSMQSILTFVPRNKNEFVSGITDLEFIQSGVPVTISP	600
		AVLPPDVTNSSFNVSPSDIVADGSMQSILTFVPRNKNEFVSGITDLEFIQSGVPVTISP	
Sbjct:	541	AVLPPDVTNSSFNVSPSDIVADGSMQSILTFVPRNKNEFVSGITDLEFIQSGVPVTISP	600
Query:	601	VTENADNYTASVVGNSVGDDITPQVGGESDLLQKRITLYPVPKITGINVNGEQFATDK	660
		VTENADNYTASVVGNSVGDDITPQVGGESDLLQKRITLYPVPKITGINVNGEQFATDK	
Sbjct:	601	VTENADNYTASVVGNSVGDDITPQVGGESDLLQKRITLYPVPKITGINVNGEQFATDK	660
Query:	661	GFPKTTFNKATFQLVMNDDVANNTQYDWTSSYAASAPVDNQGKVNIAKYTYGSTVTVTAK	720
		GFPKTTFNKATFQLVMNDDVANNTQYDWTSSYAASAPVDNQGKVNIAKYTYGSTVTVTAK	
Sbjct:	661	GFPKTTFNKATFQLVMNDDVANNTQYDWTSSYAASAPVDNQGKVNIAKYTYGSTVTVTAK	720
Query:	721	SKKFPSYATATYQFKPNLWVFSGMTSLQSSVEASRNCQRTDFTALIESARASNGSRSPDGT	780
		SKKFPSYATATYQFKPNLWVFSGMTSLOSSVEASRNCQRTDFTALIESARASNGSRSPDGT	

Sbjct: 721 SKKFPSYATATYQFKPNLWVFSGTMSLQSSVEASRNCQRTDFTALIESARASNGSRSPDGT 780

Query: 781 LWGEWGLATYDSA EWPSGNYWTKKTSTDFVTMDMTTGD IPTSAATAYPLCAEPQ 835
LWGEWGLATYDSA EWPSGNYWTKKTSTDFVTMDMTTGD IPTSAATAYPLCAEPQ

Sbjct: 781 LWGEWGLATYDSA EWPSGNYWTKKTSTDFVTMDMTTGD IPTSAATAYPLCAEPQ 835

tr Q56889 Invasin [invA] [Yersinia enterocolitica] 835 AA
align

Score = 1613 bits (4176), Expect = 0.0

Identities = 805/835 (96%), Positives = 811/835 (96%)

Query: 1 MYSFFNTLTVTKXXXXXXXXXXXXXXXXXXTYGFSQQHYFNSEALENPAEHNEAFNKIISTG 60
MYSFFNTLTVTK TYGFSQQ+YFNSEALENPAEHNEAFNKIISTG

Sbjct: 1 MYSFFNTLTVTKIISRLLSIGLIFGIFTYGFSQQNYFNSEALENPAEHNEAFNKIISTG 60

Query: 61 TSLAVSGNASNITRSMVNDAAANQEVKHWLNRFGTTQVNVNFDKKFSLKESSLDWLLPWYD 120
TSLAVSGNASNITRSMVNDAAANQEVKHWLNRFGTTQVNVNFDKKFSLKESSLDWLLPWYD

Sbjct: 61 TSLAVSGNASNITRSMVNDAAANQEVKHWLNRFGTTQVNVNFDKKFSLKESSLDWLLPWYD 120

Query: 121 SASYVFFSQLGIRNKDSRNTLNIGAGVRTFQQSWMYGFNTSYDNDMTGHNHRIGVGAEAW 180
SASYVFFSQLGIRNKDSRNTLNIGAGVRTFQQSWMYGFNT YDNDMTGHNHRIGVGAEAW

Sbjct: 121 SASYVFFSQLGIRNKDSRNTLNIGAGVRTFQQSWMYGFNTFYDNDMTGHNHRIGVGAEAW 180

Query: 181 TDYLQLSANGYFRLNGWHQSRDFADYNERPASGGDIHV KAYLPALPQLGGKLYEQYRGE 240
TDYLQLSANGYFRLNGWHQSRDF DYNERPASGGDIHV KAYLPALPQLGGKLYEQYRGE

Sbjct: 181 TDYLQLSANGYFRLNGWHQSRDFVDYNERPASGGDIHV KAYLPALPQLGGKLYEQYRGE 240

Query: 241 RVALFGKDNLQSNPYAVTTGLIYTPIPFITLGV DQRMGKSRQHEIQWNLQMDYRLGESFR 300
RVALFGKDNLQSNPYAVTTGLIYTPIPFITLGV DQRMGKSRQHEIQWNLQMDYRLGESFR

Sbjct: 241 RVALFGKDNLQSNPYAVTTGLIYTPIPFITLGV DQRMGKSRQHEIQWNLQMDYRLGESFR 300

Query: 301 SQFSPAVVAGTRLLAESRYNLVERNPNIVLEYQKQNTIKLAFSPAVLSGLPGQVYSVSAQ 360
SQFSPAVVAGTRLLAESRYNLVERNPNIVLEYQKQNTIKLAFSPAVLSGLPGQVYSVSAQ

Sbjct: 301 SQFSPAVVAGTRLLAESRYNLVERNPNIVLEYQKQNTIKLAFSPAVLSGLPGQVYSVSAQ 360

Query: 361 IQSQSALQRILW NDAQWVAAGGKLIPVSATDYNVVLPPYKMAPASRTVGKTGESEAAVN 420
IQSQSALQRILW+DAQW+A GGKLIPVSAT+YNVVLPYKMAPASRTVGKTGESEAAVN

Sbjct: 361 IQSQSALQRILWDDAQWIAVGGKLIPVSATNYNVVLPPYKMAPASRTVGKTGESEAAVN 420

Query: 421 TYTLSATAIDNHGSSNPATLTVIVQQPQFVITSEVTDDGALADGRTPITVKFTVTNIDS 480
TYTLSATAIDNHGSSNPATLTVIVQQPQFVITSEVTDDGALADGRTPITVKFTVTNID

Sbjct: 421 TYTLSATAIDNHGSSNPATLTVIVQQPQFVITSEVTDDGALADGRTPITVKFTVTNIDG 480

Query: 481 TPVAEQEGVITTSNGALPSKVTKKTD AQGVISIALTSFTVGVSVVTLDIQQQATVDVRF 540
TPVAEQEGVITTSNGALPSKVTKKTD AQGVISIALTSFTVG SVVTLDIQQQATVDVRF

Sbjct: 481 TPVAEQEGVITTSNGALPSKVTKKTD AQGVISIALTSFTVGGSVVTLDIQQQATVDVRF 540

Query: 541 AVLPPDVTNSSFNVSPSIVADGSMQSILTFVPRNKNNFVSGITDLEFIQSGVPVTISP 600
AVLPPDVTNSSFNVSPSIVADGSMQSILTFVPRNKNNFVSGITDLEFIQSGVPVTISP

Sbjct: 541 AVLPPDVTNSSFNVSPSIVADGSMQSILTFVPRNKNNFVSGITDLEFIQSGVPVTISP 600

Query: 601 VTENADNYTASVVGNSVGVDITPQVGGESLDLLQKRITLYPV PKITGINVNGEQFATDK 660
VTENADNYTASVVGNSVGVDITPQVGGESLDLLQKRITLYP+PKIT I VNGEQFATDK

Sbjct: 601 VTENADNYTASVVGNSVGVDITPQVGGESLDLLQKRITLYPIPKITSIAVNGEQFATDK 660

Query: 661 GFPKTTFNKATFQLVMNDDVANNTQYDWTSSYAASAPVDNQGKVNIAYKTYGSTVTVTAK 720
GFPKTTFNKATFQLVMNDDVANNTQYDWTSSYAASAPVDNQGKVNIAYKTYGSTVTVTAK
Sbjct: 661 GFPKTTFNKATFQLVMNDDVANNTQYDWTSSYAASAPVDNQGKVNIAYKTYGSTVTVTAK 720

Query: 721 SKKFPSYTATYQFKPNLWVFSGTMSLQSSVEASRNCQRTDFTALIESARASNGSRSPDGT 780
SKKFPSYTATYQFKPNLWVFSGTMSLQSS+EASRNCQRTDFTALIESARASNGSRSPDGT
Sbjct: 721 SKKFPSYTATYQFKPNLWVFSGTMSLQSSIEASRNCQRTDFTALIESARASNGSRSPDGT 780

Query: 781 LWGEWGLATYDSA EWPSGNYWTKKTSTDFVTMDMTTGDIPTSAATAYPLCAEPQ 835
LWGEWGLATYDSA EWPSGNYWTKKTSTDFVTMDMTTG IPTSAATAYPLCAEPQ
Sbjct: 781 LWGEWGLATYDSA EWPSGNYWTKKTSTDFVTMDMTTGAIPTSAATAYPLCAEPQ 835

sp P11922 **Invasin [Yersinia pseudotuberculosis]** 985 AA
INVA_YERPS align

Score = 726 bits (1874), Expect = 0.0

Identities = 381/695 (54%), Positives = 481/695 (68%), Gaps = 43/695 (6%)

Query: 45 NPAEHNEAFNKIISTGTSLAVSGNASNITRSMVNDAAANQEVKHWLNRFGTQVNVNFDKK 104
N + EA NK+ISTG LA SG AS++ SMV DA NQE+K WLNRFGT QVN+NFDK
Sbjct: 102 NKEQETEAVNKMISTGARLAASGRASDVASHMVGDVAVNQEIKQWLNRFGTAQVNLNFDKN 161

Query: 105 FSLKESSLDWLLPWYDSASYVFFSQLGIRNKDSRNTLNIGAGVRTFQQSWMYGFNTSYDN 164
FSLKESSLDWL PWYDSAS++FFSQLGIRNKDSRNTLN+G G+RT + W+YG NT YDN
Sbjct: 162 FSLKESSLDWLAPWYDSASFLFFSQLGIRNKDSRNTLNLGVGIRTLENGWLYGLNTFYDN 221

Query: 165 DMTGHNHRIGVGAEAWTDYQLQSANGYFRLNGWHQSRDFADYNERPASGGDIHV KAYLPA 224
D+TGHNHRIG+GAEAWTDYQLQ+ANGYFRLNGWH SRDF+DY ERPA+GGD+ AYLPA
Sbjct: 222 DLTGHNHRIGLGA EAWTDYQLAANGYFRLNGWHSSRDFSDYKERPATGGDLRANAYLPA 281

Query: 225 LPQLGGKLYEQYRGERVALFGKDNLQSNPYAVTTGLIYTPIPFITLGVDQRMGKSRQHE 284
LPQLGGKL YEQY GERVALFGKDNLQ NPYAVT G+ YTP+P +T+GVDQRMGKS +HE
Sbjct: 282 LPQLGGKLMYEQYTG ERVALFGKDNLQRNPYAVTAGINYPVPLLTGVDQRMGKSSKHE 341

Query: 285 IQWNLQMDYRLGESFRSQFSPA VVAGTRLLAESRYNLVERNPNIVLEYQKQNTIKLAFSP 344
QWNLQM+YRLGESF+SQ SP+ VAGTRLLAESRYNLV+RN NIVLEYQKQ +KL SP
Sbjct: 342 TQWNLQMN YRLGESFQSQLSPSAVAGTRLLAESRYNLVDRNNNIVLEYQKQVVKLTLSLSP 401

Query: 345 AVLSGLPGQVYSVSAQIQSQSALQRILWND AQVVAAGGKLIPVSATDYNVLPYPKMAP 404
A +SGLPGQVY V+AQ+Q SA++ I+W+DA+ +AAGG L P+S T +N+VLPPYK A
Sbjct: 402 ATISGLPGQVYQVNAQVQGASAVREIVWSDAELIAAGGTLTPLSTTQFNLVLPYPKRTAQ 461

Query: 405 ASRTVGKTGESEAAVNTYTL SATAIDNHGNSNPATLTVIVQQPQFVITSEVTDGALAD 464
SR + N Y+LSA A+D+ GN SN TL+V VQQPQ +T+ V DGA A+
Sbjct: 462 VSRVT-----DDL TANFYLSALAVDHQGNRSNSFTLSVTVQQPQLTLTAAVIGDGAPAN 516

Query: 465 GRTPITVKFTVTNIDSTPVAEQEGVITTSNGALPSKVTKKTD AQGVISIALTSFTVGVSV 524
G+T ITV+FTV + + P+A QE VITT+NGALP+K+T+KTDA GV IALT+ T GV+V
Sbjct: 517 GKTAITVEFTVADFEKGPLAQEVVITTNNGALPNKITEKTDANGVARIALTNTTGDGVTV 576

Query: 525 VTLDIQGQQATVDVRF AVLPPDVTNSSFNVSPSDIVADGSMQSILTFVPRKNKNEFVSGI 584
VT +++GQ+ +VD F S+ P+ I+ADG M S +T ++ + +G
Sbjct: 577 VTAEVEGQRQSV DTHFVKGTIAADKSTLAAVPTSIIADGLMASTITLELKDTYGD PQAGA 636

Query: 585 T---DLEFIQSGVPVTISPVTENAD-NYASVVGNSVGVDITPQVGGESLDLLQKRITL 640
 D GV +T++ D Y+A + ++G +T +V G +
 Sbjct: 637 NVAFDTTLGNMGV-----ITDHNDGTYSAPLTSTTLGVATVTVKVDGAA----- 680

Query: 641 YPVPKITGINVNGEQFATDKGFPKTTFNKATFQLVMNDDVANNTQYDWTSSYAASAPVDN 700
 + VP +T VN F D P +++F + D +A+ T SS + PVD
 Sbjct: 681 FSVPSVT---VN---FTAD---PIPDAGRSSFTVSTPDILADGTM-----SSTLSFVPVDK 727

Query: 701 QGKV-----NIAYKTYGSTVTVTAKSKKFPSYTAT 730
 G +++ G V+++ +++ SYTAT
 Sbjct: 728 NGHFIQSGMQLSFTQNGVVPVSISPITEQPDSYTAT 762

Score = 386 bits (991), Expect = e-105

Identities = 205/376 (54%), Positives = 262/376 (69%), Gaps = 10/376 (2%)

Query: 462 LADGRTPITVKFTVTNIDSTPVAEQEGVITTSNGALPSKVTKKTDACQVISIALTSFTVG 521
 +ADG T+ + + P A T+ G + +T D G S LTS T+G
 Sbjct: 612 IADGLMASTITLELKDTYGDQPAGANVAFDITLGNM-GVITDHND--GTYSAPLTSTTLG 668

Query: 522 VSVVTLDIQGGQ--ATVDVRFVAVLP-PDVTNSSFNVSFSDIVADGSMQSILTFVPRNKN 577
 V+ VT+ + G +V V F P PD SSF VS DI+ADG+M S L+FVP +KN
 Sbjct: 669 VATVTVKVDGAASFVPSVTVNFTADPIPDAGRSSFTVSTPDILADGTMSSTLSFVPVDKN 728

Query: 578 NEFVSGITDLEFIQSGVPVTISPVTENADNYASVVGNSVGVDITPQVGGESLDLLQKR 637
 F+SG+ L F Q+GVPV+ISP+TE D+YTA+VVGNSVGDV ITPQV L LQK+
 Sbjct: 729 GHFISGMQGLSFTQNGVVPVSISPITEQPDSYTATVVGNSVGVDITPQVDTLILSTLQKK 788

Query: 638 ITLYPVPKITGINVNGEQFATDKGFPKTTFNKATFQLVMNDDVANNTQYDWTSSYAASAP 697
 I+L+PVP +TGI VNG+ FATDKGFPKT F ATFQL M++DVANNTQY+W+SS+ +
 Sbjct: 789 ISLFPVPTLTGILVNGQNFATDKGFPKTIFKNATFQLQMDNDVANNTQYEWSSSFTPNVS 848

Query: 698 VDNQGVNIAYKTYGSTVTVTAKSKKFPSYTATYQFKPNLWVFSGTMSLQSSVEASRNCQ 757
 V++QG+V I Y+TY S V VTAKSKKFPSY+ +Y+F PN W++ G SL SS+EASR CQ
 Sbjct: 849 VNDQGQVTITYQTY-SEVAVTAKSKKFPSYSVSRYFYPNRWIYDGGRLVSSLEASRQCQ 907

Query: 758 RTDFTALIESARASNGSRSPDGTLLWGEWGSLATYDSAEPWPSGNYWTKKTSTDFVTMDMTT 817
 +D +A++ES+RA+NG+R+PDGTLLWGEWGS Y S++W SG YW KKTSTDF TM+M T
 Sbjct: 908 GSDMSAVLESSRATNGTRAPDGTLLWGEWGS TAY-SSDWQSGEYVWKKTSTDFETMNMDT 966

Query: 818 GDI-PTSAATAYPLCA 832
 G + P A A+PLCA
 Sbjct: 967 GALQPGPAYLAFPLCA 982

tr Q56937 Invasin [inv] [Yersinia pestis] 987 AA
 align

Score = 721 bits (1861), Expect = 0.0

Identities = 381/696 (54%), Positives = 480/696 (68%), Gaps = 44/696 (6%)

Query: 45 NPAEHNEAFNKIISTGTSLAVSGNASNITRSMVNDAAEQEVKHWLNRFGTTQVNVNFDKK 104
 N + EA NK+ISTG LA SG AS++ SMV DA NQE+K WLNRFGT QVN+NFDK
 Sbjct: 103 NKEQETEAVNKMISTGARLAASGRASDVASHMVGDAVNQEIKQWLNRFGTQVNLNFDKN 162

Query: 105 FSLKESSLDWLLPWYDSASYVFFSQLGIRNKDSRNTLNIGAGVRTFQQSWMYGFNTSYDN 164

FSLKESSLDWL PWYDSAS++FFSQLGIRNKDSRNTLN+G G+RT + W+YG NT YDN
 Sbjct: 163 FSLKESSLDWLAPWYDSASFLFFSQLGIRNKDSRNTLN+G VGTLENGWLYGLNTFYDN 222
 Query: 165 DMTGHNHRIGVGAEAWTDYLQLSANGYFRLNGWHQSRDFADYNERPASGGDIHVKAYLPA 224
 D+TGHNHRIG+GAEAWTDYLQL+ANGYFRLNGWH SRDF+DY ERPA+GGD+ AYLPA
 Sbjct: 223 DLTGHNHRIGLGAETDYLQLAANGYFRLNGWHSSRDFSDYKERPATGGDLRANAYLPA 282
 Query: 225 LPQLGGKLYEQYRGERVA-LFGKDNLQSNPYAVTTGLIYTPFITLGVQDQRMGKSRQH 283
 LPQLGGKL YEQY GERV LFGKDNLQ NPYAVT G+ YTP+P +T+GVDQRMGKS +H
 Sbjct: 283 LPQLGGKLMYEQYTGERVALLFGKDNLQRPYAVTAGINYPVPLLTGVQDQRMGKSSKH 342
 Query: 284 BIQWNLMQMDYRLGESFRSQSPAVVAGTRLLAESRYNLVERNPNIVLEYQKQNTIKLAFS 343
 E QWNLMQ+YRLGESF+SQ SP+ VAGTRLLAESRYNLV+RN NIVLEYQKQ +KL S
 Sbjct: 343 ETQWNLMQNYRLGESFQSQLSPSAVAGTRLLAESRYNLVDRNNIVLEYQKQVVKLTLS 402
 Query: 344 PAVLSGLPGQVYSVSAQIQSQSALQIRILWDAQWVAAGGKLIPVSATDYNVVLPPYKPA 403
 PA +SGLPGQVY V+AQ+Q SA++ I+W+DA+ +AAGG L P+S T +N+VLPPYK A
 Sbjct: 403 PATISGLPGQVYQVNAQVQASAVREIVWSDAELIAAGGTLTPLSTTQFNLVLPPYKRTA 462
 Query: 404 PASRTVGKTGESEAAVNTYTLSTATAIDNHGNSNPATLTVIVQQPQFVITSEVTDDGALA 463
 SR + N Y+LSA A+D+ GN SN TL+V VQQPQ +T+ V DGA A
 Sbjct: 463 QVSRVT-----DDLTA NFYSLSALAVDHQGNRSNSFTLSVTVQQPQLTLTA AVIGDGAPA 517
 Query: 464 DGRTPIITVKFTVTNIDSTPVAEQEGVITTSNGALPSKVTKKTDAGQVISIALTSFTVGVS 523
 G+T ITV+FTV + + P+A QE VITT+NGALP+K+T+KTDA GV IALT+ T GV+
 Sbjct: 518 SGKTAITVEFTVADFEKGKPLAGQEVVITTNNGALPNKITEKTDANGVARIALTNTTGDVT 577
 Query: 524 VVTLDIQGQQATVDVRFVAVLPDVTNSSFVNSPSDIVADGSMQSILTFVPRNKNNEFVSG 583
 VVT +++GQ+ +VD F S+ P+ I+ADG M S +T ++ + +G
 Sbjct: 578 VVTAEEVGQRQSVDTFVKGITIAADKSTLAAVPTSIIADGLMASTITLELKDTYGDPOAG 637
 Query: 584 IT---DLEFIQSGVPVTISPVTENAD-NYASVVGNSVGVDVITPQVGESLDLLQKRIT 639
 D GV +T++ D Y+A + ++G +T +V G +
 Sbjct: 638 ANVAFDITLGNMGV-----ITDHNDGTYSAPLTSTTLGVATVTVKVDGAA----- 682
 Query: 640 LYPVPKITGINVNGEQFATDKGFPKTTFNKATFQLVMNDDVANNTQYDWTSSYAASAPVD 699
 + VP +T VN F D P +++F + D +A+ T SS + PVD
 Sbjct: 683 -FSVPSVT---VN---FTAD---PIPDAGRSSFTVSTPDILADGTM---SSTLSFVPVD 728
 Query: 700 NQGVK-----NIAYKTYGSTVTVTAKSKKFPSYTAT 730
 G +++ G V+++ +++ SYTAT
 Sbjct: 729 KNGHFISGMQGLSFTQNGVPSISPITEQPDSYTAT 764

Score = 380 bits (975), Expect = e-104

Identities = 201/376 (53%), Positives = 260/376 (68%), Gaps = 10/376 (2%)

Query: 462 LADGRTPIITVKFTVTNIDSTPVAEQEGVITTSNGALPSKVTKKTDAGQVISIALTSFTVG 521
 +ADG T+ + + P A T+ G + +T D G S LTS T+G
 Sbjct: 614 IADGLMASTITLELKDTYGDPOAGANVAFDITLGNM-GVITDHND--GTYSAPLTSTTLG 670
 Query: 522 VSVVTLDIQGQQ--ATVDVRFVAVLP-PDVTNSSFVNSPSDIVADGSMQSILTFVPRNKN 577
 V+ VT+ + G +V V F P PD SSF VS DI+ADG+M S L+FVP +KN
 Sbjct: 671 VATVTVKVDGAASFVPSVTVNFTADPIPDAGRSSFTVSTPDILADGTMSSTLSFVPVDKN 730
 Query: 578 NEFVSGITDLEFIQSGVPVTISPVTENADNYASVVGNSVGVDVITPQVGESLDLLQKR 637
 F+SG+ L F Q+GVPV+ISP+TE D+YTA+VVG+N+ GDV ITP V L LQK+
 Sbjct: 731 GHFISGMQGLSFTQNGVPSISPITEQPDSYTATVVGNTAGDVTITPLVDTLILSTLQKK 790

Query: 638 ITLYPVPKITGINVNGEQFATDKGFPKTTFNKATFQLVMNDDVANNTQYDWTSSYAASAP 697
I+L+PVP +TGI VNG+ FATDKGFPKT F ATFQL M++DVANNTQY+W+SS+ +
Sbjct: 791 ISLFPVPTLTGILVNGQNFATDKGFPKTIFKNATFQLQMDNDVANNTQYEWSSSFTPNVS 850

Query: 698 VDNQGKVNIAKYTGSTVTVTAKSKKFPSYATYQFKPNLWVFSGTMSLQSSVEASRNCQ 757
V++QG+V I Y+TY S V VTAKSKKFPSY+ +Y+F PN W++ G SL SS+EASR CQ
Sbjct: 851 VNDQGQVTITYQTY-SEVAVTAKSKKFPSYSVSYRFYPNRWIYDGGTSLVSSIEASRQCQ 909

Query: 758 RTDFTALIESARASNGSRSPDGTWGEWGSLATYDSAEWPSGNYWTKKTSTDFVTMDMTT 817
+D +A++ES+RA+NG+R+PDGTWGEWGS L Y S++W SG YW K+TSTDF TM+M T
Sbjct: 910 GSDMSAVLESSRATNGTRAPDGTWGEWGS LTAY-SSDWQSGEYWVKRTSTDFETNMNMT 968

Query: 818 GDI-PTSAATAYPLCA 832
G + P A A+PLCA
Sbjct: 969 GLLQPGPAYLAFFPLCA 984

tr Q8ZA73 Putative invasin (Hypothetical) [YPO3944] [Yersinia pestis] 3013 AA

align

Score = 428 bits (1100), Expect = e-118
Identities = 265/703 (37%), Positives = 382/703 (53%), Gaps = 38/703 (5%)

Query: 66 SGNASNITRSMVNDAAANQEVKHWLNRFGTQVNVNFDKKFSLKESSLDWLLPWYDSASYV 125
S ASN+ RS V + N + WLN+FGT +V +N D F L S+LD L+P DS S +
Sbjct: 178 SDAASNMARSAVTNEINASSQQWLNQFGTARVQLNVDSDFKLDNSALDLLVPLKDSSESL 237

Query: 126 FFSQLGIRNKDSRNTLNIGAGVRTFQQSWMYGFNTSYDNDMTGHNHRIGVGAEAWTDYLQ 185
F+QLG+RNKDSRNT+NIGAG+R +Q WMYG NT +DND+TG N R+GVGAE TDYL+
Sbjct: 238 LFTQLGVRNKDSRNTVNIGAGIRQYQGDWYGANTFFDNDLTGKNRRVGVGAEVATDYLK 297

Query: 186 LSANGYFRLNGWHQSRDFADYNERPASGGDIHVKAYLPALPQLGGKLYEQYRGERVALF 245
SAN YF L GWHQSRDF+ Y+ERPA G DI +AYLPA PQLGGKL YE+YRG+ VALF
Sbjct: 298 FSANTYFGLTGWHQSRDFSSYDERPADGFDIRTEAYLPAYPQLGGKLYEYRGDEVALF 357

Query: 246 GKDNLQSNPYAVTTGLIYTPIPFITLGVDQRMGKSRQHEIQWNLQMDYRLGESFRSQFSP 305
GKD+ Q +P+AVT G+ YTP+P +T+G + R GK + N+Q++YR+G+ + Q
Sbjct: 358 GKDDRQKDPHAVTLGVNYTPVPLVTIGAEHREGKGNNTSVNVQLNYRMGQPWNDQIDQ 417

Query: 306 AVVAGTRLLAESRYNLVERNPNIVLEYQKQNTIKLAFSPA VLSGLPGQVYSVSAQIQSQS 365
+ VA R LA SRY+LVERN NIVL+Y+KQ I L P +SG G +++AQ++++
Sbjct: 418 SAVAANRTLASRYDLVERN NIVLDYKKQELIHLVL-PDRISGSGGGAITLTAQVRAKY 476

Query: 366 ALQRIWLNDAQWVAAGGKLIPVSATDYNVVLPPYKPMAPASRTVGKTGESEAAVNTYTLS 425
RI W+ AGG P++ + +V LP Y+ + S NT+T+S
Sbjct: 477 GFSRIEWDATPLENAGGSTSPLTQSSLSVTLFPYQHILRTS-----NTHTIS 523

Query: 426 ATAIDNHGSSNPATLTIVIQQPQFVITSEV--TDDGALADGRTPITVKFTVTNIDSTPV 483
A A D GN+SN A ++ V +P+ ++ S + T D A A+G TV+ TVT+ D P+
Sbjct: 524 AVAYDAQGNASNRVTSIEVTRPETMVISHLATTIDNATANGIATNTVQATVTDGDGQPI 583

Query: 484 AEQEGVITTSNGALPSKVTKKTDAGGVISIALTSFTVGVS VVTLDIQQQATVDVRFVAV- 542
Q + A S +T A G S LT GVS V++ + +VD F
Sbjct: 584 IGQLINFAVNTQATLSTTEARTGANGTASTTLTHTVSGVSRVSVTLGSSSRSDTTTFVAD 643

Query: 543 -LPPDVTNSSFNVSPSDIVADGSMQSILTFVPRNKNNEFVSGITDLEFIQSGVPVTISP 601
 ++T ++ V+ +D VA+GS +++ + V+ + + +G V V
 Sbjct: 644 ESTAEITAANLTVTTNDSVANGSDTNVVRKVTDAYTNAVANQSVIFSASNGATVIDQTV 703

Query: 602 TENADNYTASVVGNSVGDVD-ITPQVGESLDLLQKRITLYPVPKITGINVNGEQFATDK 660
 NA+ S + N+ V +T +GG+S Q T P V A
 Sbjct: 704 ITNAEGIADSTLTNTTAGVSVVTATLGGQS---QQVDITFKPGSTAASLVKLADRAVAD 760

Query: 661 GFPKTTFNKATFQLVMNDDVAN-----NTQYDWTSSYAASAP-VDNQGKVNIAY---K 709
 G ++ Q+V+ D N + Q D + AS P G +N + +
 Sbjct: 761 G-----IDQNEIQVVLRDGTGNAVPNVPMISIQADNGAIVVASTPNTGVDGTINATFTNLR 815

Query: 710 TYGSTVTVTAKSKKFPSYTATYQFKPNLWVFSGTMSLQSSVEA 752
 S V+VT+ + + T T+ P V S ++ ++ +A
 Sbjct: 816 AGESVSVSTSPALVGMTMTMTFSADPRTAVVSTLAAIDNNAK 858

Score = 122 bits (305), Expect = 4e-26

Identities = 123/477 (25%), Positives = 199/477 (40%), Gaps = 28/477 (5%)

Query: 374 DAQWVAAGGKLIPVSATDYNVVLPPYKMPAPASRTVGKTGESEAAVNTYTLSATAIDNHG 433
 DA GG+ + SAT+ + ++ +V + S A + ++ T + N+
 Sbjct: 2548 DANQNPVGGQQVAFSATNEVTLTESNGSISTPEGSVLLSVTSTQA-GVHPITGTLVSNYY 2606

Query: 434 NSSNPATLTVIVQQPQFVITSEVTDGALADGRTPITVKFTVTNIDSTPVAEQEGVITTS 493
 + A Q + T V D+ ALADG T V+ V + VA+ T +
 Sbjct: 2607 TDTFGAAFIANKNTAQ-LSTLMVVDNNALADGVTRNQVRAHVVDSTGNSVADMAVTF TAN 2665

Query: 494 NGALPSKVTKKTDAQGVISIALTSFTVGVSVVT--LDIQGQATVDVRFVAVLPPDVTNSS 551
 GA SKVT TD G LT+ VGV+VVT L G TVD F P + +
 Sbjct: 2666 RGAQLSKVTVLTDNNGDAVNTLTNSLVGVTVVTA KLGTAGTPLTVDTVFTAGP--LATLT 2723

Query: 552 FNVSPSDIVADGSMQSILTFVPRNKNNEFVSGITDLEFIQSGVPVTISP-VTENADNYTA 610
 + ++ AD S + + ++ + + G +G +T + NA+
 Sbjct: 2724 LVTTVNNAFADNSATNTVQATLKDVS GNPIVGEVVAFAASNGATITATDGGVSNANGIVL 2783

Query: 611 SVVGNSVGDVDITPQVGESLDLLQKRITLYPVPKITGINVNGEQFATDKGFPKTTFNKA 670
 + + N V V L + T + K + VNG F D GFP T F A
 Sbjct: 2784 ATLTNGTAGVS---TVTATIETLTETDTTFIAMKNLDVTVNGTTFNGDAGFPTTGFGVA 2840

Query: 671 TFQLVMNDDVANNTQYDWTSSYAASAPVDNQKV--NIAYKTYGSTVTVTAKSKKFPSYT 728
 TF++ D N+ YDW+SS A V G V N + T T+T++A K S
 Sbjct: 2841 TFKVNSGGD---NSLYDWSSSAPALVSVSGDGVVTFNAVFPPTGTPTITISATPKGGGS-P 2896

Query: 729 ATYQFKPNLWVFSGTMSLQSSVEASRNCQRTDF-----TALIESARASNGSRSPDGLWG 783
 +Y F+ N W + + + +A +C+ + T + +A +G R+ G LW
 Sbjct: 2897 LSYFVRNQWFINNNGATLNRADAITHCENVGYTMTPTSTQVTNAATWMSGKRAV-GNLWS 2955

Query: 784 EWGSLATYDSA EW-PSGNYWTKKT--STDFVTMDMTTGDIPTSA--ATAYPLCAEP 834
 EWG + Y + W P+ +W ++ + + ++TG + T A + +C P
 Sbjct: 2956 EWGDFSAYTAPGWVPAEFFWLSNNHDASTALAI GLSTGTLTMTMGDFMAITHVMCTRP 3012

Score = 60.1 bits (144), Expect = 2e-07

Identities = 54/203 (26%), Positives = 86/203 (41%), Gaps = 7/203 (3%)

Query: 418 AVNTYTLSATAIDNH--GNSSNPATLTVIVQQPQFVITSE--VTDDGALADGRTPITVK 472

A T+TL+ T++ GN T I + I + +T + ALADG +
 Sbjct: 1983 ATLTHTLAGTSVVTARVGNRVQSKDTTFIADRTTATIRASDLTITRNNALADGVATNAAR 2042
 Query: 473 FTVTNIDSTPVAEQEGVITTSNGALPSKVTKKTDAGQVISIALTSFTVGVSVVTLDI--Q 530
 VT+ + PV T+ NGAL + + TD+ G S T T G+S VT I
 Sbjct: 2043 VIVTDANGNPVPSMFVGYTSDNGALLTPTSGMTDSSGTFSTTFTHTTAGISKVTAIVTM 2102
 Query: 531 GQQATVDVRFVAVLPPDVTNSSFNVSPSDIVADGSMQSILTFVPRNKNEFVSGITDLEFI 590
 G T D F S V +D +A+ S ++I+ ++ + V+G+
 Sbjct: 2103 GISQTKDAVFIADRSTAHVSELIVVKNDLANNSDRNIVQAHIKDAHGNVVTGMNVNFS 2162
 Query: 591 QSGVPVTISPVTENADNYTASVV 613
 V +T + VT N+ Y + +
 Sbjct: 2163 TENVTLTANTVTTNSQGYAENTL 2185

Score = 53.5 bits (127), Expect = 2e-05

Identities = 62/250 (24%), Positives = 93/250 (36%), Gaps = 10/250 (4%)

Query: 373 NDAQVWAAGGKLIPVSATDYNVVLPPYKPMAPASRTVGKTGESEAAVNTYTLSATAIDNH 432
 N Q V + PV T VV A + +G TG A T T + N
 Sbjct: 962 NSVQAVVSDSGGNPV--TGATVVFSSTNATAQVTTVIGTTGVDGIATATLTNTVAGTSNV 1019
 Query: 433 G---NSSNPATLTIVVQQPQFVITSEVTDDGALADGRTPITVKFTVTNIDSTPVAEQEGV 489
 ++ N T V IT +GA+ADG V V + + P+ V
 Sbjct: 1020 VATIDTVNANIDTAFVAGAVATITLTAPVNGAVADGADTNQVDALVEDANGNPITGAADV 1079
 Query: 490 ITTSNGALPSKVTKKTDAGQVISIALTSFTVGVSVVTLDIQGGQATVDVRFVAVLPPDVTN 549
 +++NGA T T GV S LT G S V + A +D F + V
 Sbjct: 1080 FSSANGATILSSTMTGVNGVASTLLTHTVAGTSNVVATVDTVNANIDTTF--VAGAVAT 1137
 Query: 550 SSFNVSPSDIVADGSMQSILTFVPRNKNEFVSGITDLEFIQSGVPVTISPV--TENADN 607
 + + VADG+ + + V + + V+G + F + I+ V T AD
 Sbjct: 1138 ITLTPVNGAVADGANSNSVQAVVSDSDGNPVTGAA-VVFSSANATAQITTVIGTTGADG 1196
 Query: 608 YTASVVGNSV 617
 + + N+V
 Sbjct: 1197 IATATLTNTV 1206

Score = 51.6 bits (122), Expect = 6e-05

Identities = 63/269 (23%), Positives = 108/269 (39%), Gaps = 30/269 (11%)

Query: 458 DDGALADGRTPITVKFTVTNIDSTPVAEQEGVITTSNGALPSKVTKKTDAGQVISIALTS 517
 +DGA+ADG V+ V ++ V V + NG + +TDA G + +++
 Sbjct: 2226 NDGAVADGIQTNRVEARVYDVSDNLVPNSNVVFSADNGGQLVQNDVQTDALGSAYVTVS 2285
 Query: 518 FTVGVSVVTLDIQGGQATVDVRFVAVLPPDVT--NSSFNVSPSDIVADGSMQSILTFVPRN 575
 GV+ VT+ G A+ F T F ++ + VA+G +++ + +
 Sbjct: 2286 INTGVTKVTVTADGVSASTTTTFIADRTATLTDRFLITHDNAVANGVVENRVLHLVD 2345
 Query: 576 KNEFVSGITDLEFIQSGVPVTISPVTENADNYTASVVGNSV-GDVDITPQVGESLDLL 634
 N+ VSG+ +G + S +T+ + + V+ N++ G D+T
 Sbjct: 2346 ANDNSVSGVEVNFSAITNGASINASAITD-INGFAIGVLTNTLSGSPSDVT----- 2393
 Query: 635 QKRITLYPVPKITGINVNGEQFATDKGFPKTTFNKATFQLVMNDDVANNTQYDWTSSYAA 694
 +TL + V QF D T N AT V+ DD A D ++
 Sbjct: 2394 ---VTLVTPGGTESLTVT-PQFIAD----INTANIATGDFVIIDDGAVANSVD--ANEVR 2443

Query: 695 SAPVDNQGKVNIAKY-----GSTVTVT 718
 + DNQG Y G+T+T +
 Sbjct: 2444 ARVTDNQGNIAIGYSVVFSSQNGATITTS 2472

Score = 48.5 bits (114), Expect = 5e-04
 Identities = 72/292 (24%), Positives = 101/292 (33%), Gaps = 39/292 (13%)

Query: 373 NDAQWVAAGGKLIPVSATDYNVVLPPYKPMAPASRTVGKTGESEAAVNTYTLSATAIDNH 432
 N Q V + PV T VV A + +G TG A T T + N
 Sbjct: 1155 NSVQAVVSDSDGNPV--TGAAVVFSSANATAQITTVIGTTGADGIATATLTNTVAGTSNV 1212

Query: 433 G---NSSNPATLTVIVQQPQFVITSEVTDDGALADGRTPITVKFTVTNIDSTPVAEQEGV 489
 ++ N T V IT +GA+ADG V V + + + V
 Sbjct: 1213 VATIDTVNANIDTAFVAGAVATITLTAPVNGAVADGADTNQVDALVQDANGNAITGAADV 1272

Query: 490 ITTNGALPSKVTKKTDAGQVISIALTSFTVGVSVVTLDIQGGQATVDVRFVLPDPVTN 549
 +++NGA T T GV S LT G S V I A +D F + V
 Sbjct: 1273 FSSANGADIIAPTMTGTVNGVASTLLTHTVAGTSNVVATIDTISANIDTAF--VAGAVAT 1330

Query: 550 SSFNVSPSDIVADGSMQSILTFVPRKNNEFVSG----- 583
 + + VADG+ + + + N ++G
 Sbjct: 1331 ITLTAPVNGAVADGADTNQVDALVEDANGNPITGAAVVFSSANGATILSSTMNTGVNGVA 1390

Query: 584 ---ITDLEFIQSGVPVTISPVTENADNYTASVVGNSVGDVDITPQVGESLD 632
 +T S V TI VTEN D TA V G +V + +T V G D
 Sbjct: 1391 STFLTHTVAGTSNVVATIGSVTENID--TAFVAG-AVATITLTAPVNGAVAD 1439

Score = 47.0 bits (110), Expect = 0.001
 Identities = 46/178 (25%), Positives = 69/178 (37%), Gaps = 5/178 (2%)

Query: 442 TVIVQQPQFVITSEVTDDGALADGRTPITVKFTVTNIDSTPVAEQEGVITTSNGALPSKV 501
 T V IT +GA+ADG V V + + P+ V +++NGA
 Sbjct: 1321 TAFVAGAVATITLTAPVNGAVADGADTNQVDALVEDANGNPITGAAVVFSSANGATILSS 1380

Query: 502 TKKTDAGQVISIALTSFTVGVSVVTLDIQGGQATVDVRFVLPDPVTNSSFNVSPSDIVA 561
 T T GV S LT G S V I +D F + V + + VA
 Sbjct: 1381 TMNTGVNGVASTFLTHTVAGTSNVVATIGSVTENIDTAF--VAGAVATITLTAPVNGAVA 1438

Query: 562 DGSMQSILTFVPRKNNEFVSGITDLEFIQSGVPVTISPV--TENADNYTASVVGNSV 617
 DG + + V + + V+G T + F + I+ V T AD + + N+V
 Sbjct: 1439 DGVNTNSVQAVVSDSDGNAVGTAT-VVFSSANATAQITTVIGTTGADGIATATLTNTV 1495

Score = 45.1 bits (105), Expect = 0.006
 Identities = 52/210 (24%), Positives = 90/210 (42%), Gaps = 15/210 (7%)

Query: 425 SATAIDNHGSSNPATLTVI--VQQPQFVITSEVTDDGALADGRTPITVKFTVTNIDSTP 482
 S T I G A +T + V Q +TS + D LA+G+ + VT+
 Sbjct: 1595 SNTVIAISGAHQGYARVTFVADVSTAQLKLTSLF--DNQLANGKAGNIAQALVTD AHDNL 1652

Query: 483 VAEQEGVITTSNGALPSKVTKKTDAGQVISIALTSFTVGVSVV--TLDIQGGQATVDVRF 540
 +A Q NGA+ + A G++ + + G++ V TLD GQ T++ F
 Sbjct: 1653 LANQSVSFALDNGAVIESQGDASSAGIVLMRFNNTLAGMTTVTATLDSTGQTETLETHF 1712

Query: 541 A---VLPPDVTNSSFNVSPSDIVADGSMQSILTFVPRKNNEFVSGITDLEFIQSGVPVT 597

++T + N ++I + +Q ++T V N N V +T SG+ +T
 Sbjct: 1713 VAGKAASIEMTMKDNAVANNIDTN-EVQVLVTDVDGNAINGAVVNLTS---NSGMNIT 1767

Query: 598 ISPV TENAD-NYTASVVGNSVGDVDITPQV 626

+ VT +D TA++ G + I ++

Sbjct: 1768 PNSVTTGSDGTATATLTHTLAGSLPINARI 1797

Score = 43.9 bits (102), Expect = 0.013

Identities = 34/114 (29%), Positives = 51/114 (43%), Gaps = 4/114 (3%)

Query: 456 VTDDGALADGRTPITVKFTVTNIDSTPVAEQEGVITTSNGALPSKVTKKTDAGQGVISIAL 515

+ DDGA+A+ V+ VT+ +A V ++ NGA + + T G S L

Sbjct: 2427 IIDDGAVANSVDANEVRARVTDNQGNAIAGYSVVFSSQNGATIT-TSGITGVDGWASAKL 2485

Query: 516 TSFTVGVS VVTLDIQQQQATVDVRFVLPDVTNSS---FNVSPSDIVADGSMQ 566

T G S + + ATV DV+ ++ FN +P I+ADG MQ

Sbjct: 2486 THIKAGESGILARLSRPMATVHTLMPYFIADVSTATLQLFNFNPIPIIADGVMQ 2539

Score = 43.9 bits (102), Expect = 0.013

Identities = 59/303 (19%), Positives = 120/303 (39%), Gaps = 28/303 (9%)

Query: 424 LSATAIDNHGSSNPATLTVIVQQPQFVITSE--VTDDGALADGRTPITVKFTVTNIDST 481

++AT + + G+S N T + + S+ V D A+AD V VT+

Sbjct: 1893 VTATVSS-GSSRNIDTTFIADVTTAHIAASDLMVIVDDAVADNLDKNEVHARVTDAGN 1951

Query: 482 PVAEQEGVITTSNGALPSKVTKKTDAGQGVISIALTSFTVGVS VVTLDIQQQQATVDVRF 541

++ Q + T+ NGA + V +D G+ LT G SVVT + + + D F

Sbjct: 1952 VLSGQTVIFTSNGAAITTVNGISDGDGLTKATLTHTLAGTSVVTARVGNRVQSKDTTFI 2011

Query: 542 V--LPPDVTNSSFNVSPSDIVADGSMQSILTFVPRNKNNEFVSGITDLEFIQSGVPVT-I 598

+ S ++ ++ +ADG + + + N V + +G +T

Sbjct: 2012 ADRTTATIRASDLTITRNALADGVATNAARVIVTDANGNPVPSMFVGYTSDNGALLTPT 2071

Query: 599 SPVTENADNYTASVVGNSVGDVDITPQVGGESLDLLQKRITLYPVPKITGINVNGEQFAT 658

S +T+++ ++ + + G +T + + + + F

Sbjct: 2072 SGM TDSSGTFSTTFHTHTAGISKVTAAIVTMGISQTKDAV-----FIA 2114

Query: 659 DKGFPKTTFNKATFQLVMNDDVANNTQYDWTSSYAASAPVDNQGKVNIAYKTYGSTVTVT 718

D ++T + + +V ND +ANN+ + ++ A + +N+ + VT+T

Sbjct: 2115 D----RSTAHVSELIVVKNDLANNSDRNIVQAHIKDAHGNVVTGMNVNFSA-TENVTLT 2169

Query: 719 AKS 721

A +

Sbjct: 2170 ANT 2172

Score = 39.3 bits (90), Expect = 0.31

Identities = 34/158 (21%), Positives = 64/158 (39%), Gaps = 1/158 (0%)

Query: 459 DGALADGRTPITVKFTVTNIDSTPVAEQEGVITTSNGALPSKVTKKTDAGQGVISIALTSF 518

D A+ADG ++ + + V I NGA+ T T G I+ T+

Sbjct: 755 DRAVADGIDQNEIQVVL RDGTGNAV PNPMSIQADNGAIVVASTPNTGVDGTINATFTNL 814

Query: 519 TVGVSVVTLDIQQQQA-TVDVRFVLPDVTNSSFNVSPSDIVADGSMQSILTFVPRNKN 577

G SVV++ T+ + F+ P S+ ++ ADG+ +++ + N

Sbjct: 815 RAGESVSVTSPALVGMTMTMTFSADPRTAVVSTLAAIDNNAKADGTD TNVVRWVVDAN 874

Query: 578 NEFVSGITDLEFIQSGVPVTISPVTENADNYTASVVG 615
 V G++ +G + +PV + + Y + + N
 Sbjct: 875 GNSVPGVSVTFDAGNGAVLAQNPVVTDRNGYAENTLTN 912

tr Q74QZ6 Putative invasin [YP3306] [Yersinia pestis] 3108 AA
align

Score = 428 bits (1100), Expect = e-118
 Identities = 265/703 (37%), Positives = 382/703 (53%), Gaps = 38/703 (5%)

Query: 66 SGNASNITRSMVNDAAANQEVKHWLNRFGTQVNVNFDKKFSLKESSLDWLLPWYDSASYV 125
 S ASN+ RS V + N + WLN+FGT +V +N D F L S+LD L+P DS S +
 Sbjct: 178 SDAASNMARSAVTNEINASSQQWLNQFGTARVQLNVDSDFKLDNSALDLLVPLKDSSESL 237

Query: 126 FFSQLGIRNKDSRNTLNIGAGVRTFQQSWMYGFNTSYDNDMTGHNHRIGVGAEAWTDYLQ 185
 F+QLG+RNKDSRNT+NIGAG+R +Q WMYG NT +DND+TG N R+GVGAE TDYL+
 Sbjct: 238 LFTQLGVRNKDSRNTVNIGAGIRQYQGDWMYGANTFFDNDLTGKNRRRVGVGAEVATDYLK 297

Query: 186 LSANGYFRLNGWHQSRDFADYNERPASGGDIHVKAYLPALPQLGGKLYEQYRGERVALF 245
 SAN YF L GWHQSRDF+ Y+ERPA G DI +AYLPA PQLGGKL YE+YRG+ VALF
 Sbjct: 298 FSANTYFGLTGWHQSRDFSSYDERPADGFDIRTEAYLPAYPQLGGKLMYEKYRGDEVALF 357

Query: 246 GKDNLQSNPYAVTTGLIYTPIPFITLGVDQRMGKSRQHEIQWNLQMDYRLGESFRSQFSP 305
 GKD+ Q +P+AVT G+ YTP+P +T+G + R GK + N+Q++YR+G+ + Q
 Sbjct: 358 GKDDRQKDPHAVTLGVNYTPVPLVTIGAEHREGKGNNNNTSVNVQLNYRMGQPWNDQIDQ 417

Query: 306 AVVAGTRLLAESRYNLVERNPNIVLEYQKQNTIKLAFSPAVALSGLPGQVYSVSAQIQSQS 365
 + VA R LA SRY+LVERN NIVL+Y+KQ I L P +SG G +++AQ++++
 Sbjct: 418 SAVAANRTLAGSRYDLVERNINNIVLDYKKQELIHLVL-PDRISGSGGGAITLTAQVRACY 476

Query: 366 ALQRILWNDAAQWVAAGGKLIPVSATDYNVVLPPYKPMAPASRTVGKTGESEAAVNTYTLS 425
 RI W+ AGG P++ + +V LP Y+ + S NT+T+S
 Sbjct: 477 GFSRIEWDATPLENAGGSTSPLTQSSLSVTLFPYQHILRTS-----NTHTIS 523

Query: 426 ATAIDNHGSSNPATLTVIVQQPQFVITSEV--TDDGALADGRTPITVKFTVTNIDSTPV 483
 A A D GN+SN A ++ V +P+ ++ S + T D A A+G TV+ TVT+ D P+
 Sbjct: 524 AVAYDAQGNASNAVTSIEVTRPETMVISHLATIDNATANGIATNTVQATVTDGDGQPI 583

Query: 484 AEQEGVITTSNGALPSKVTKKTDAGQVISIALTSFTVGVSVVTLDIQQQATVDVRFVAV- 542
 Q + A S +T A G S LT GVS V++ + +VD F
 Sbjct: 584 IGQLINFAVNTQATLSTTEARTGANGTASTTLTHTVSGVSRVSVTLGSSSRSDTTFVAD 643

Query: 543 -LPPDVTNSSFNVSPSDIVADGSMQSILTFVPRNKNNEFVSGITDLEFIQSGVPVTISP 601
 ++T ++ V+ +D VA+GS +++ + V+ + + +G V V
 Sbjct: 644 ESTAEITAANLTVTTNDSVANGSDTNVVRKVTDAYTNAVANQSVIFSASNGATVIDQTV 703

Query: 602 TENADNYTASVVGNSVGDVD-ITPQVGGESLDLLQKRITLYPVPKITGINVNGEQFATDK 660
 NA+ S + N+ V +T +GG+S Q T P V A
 Sbjct: 704 ITNAEGIADSTLTNTTAGVSVVTATLGGQS---QQVDTTFKPGSTAAISLVKLADRAVAD 760

Query: 661 GFPKTTFNKATFQLVMNDDVAN-----NTQYDWTSSYAASAP-VDNQGKVNIAY---K 709
 G ++ Q+V+ D N + Q D + AS P G +N + +
 Sbjct: 761 G-----IDQNEIQVVLRDGTGNAVNPVMSIQADNGAIVVASTPNTGVDGTINATFTNLR 815

Query: 710 TYGSTVTVTAKSKKFPSYATYQFKPNLWVFSGTMSLQSSVEA 752
S V+VT+ + + T T+ P V S ++ ++ +A
Sbjct: 816 AGESVSVTSPALVGMTMTMTFSADPRTAVVSTLAAIDNNAKA 858

Score = 122 bits (305), Expect = 4e-26
Identities = 123/477 (25%), Positives = 199/477 (40%), Gaps = 28/477 (5%)

Query: 374 DAQWVAAGGKLIPVSATDYNVVLPPYKPMAPASRTVGKTGESEAAVNTYTLSATAIDNHG 433
DA GG+ + SAT+ + ++ +V + S A + ++ T + N+
Sbjct: 2643 DANQNPVGGQQVAFSATNEVTLTESNGSISTPEGSVLLSVTSTQA-GVHPITGTLVSNNY 2701

Query: 434 NSSNPATLTVIVQQPQFVITSEVTDGALADGRTPITVKFTVTNIDSTPVAEQEGVITTS 493
+ A Q + T V D+ ALADG T V+ V + VA+ T +
Sbjct: 2702 TDTFGAAFIANKNTAQ-LSTLMVVDNNALADGVTRNQVRAHVVDSTGNSVADMAVTFITAN 2760

Query: 494 NGALPSKVTKKTDAGQVISIALTSFTVGVSVVT--LDIQGQQATVDVRFVAVLPPDVTNSS 551
GA SKVT TD G LT+ VGV+VVT L G TVD F P + +
Sbjct: 2761 RGAQLSKVTVLTDNNGDAVNTLTNSLVGVTVVTAKLGTAGTPLTVDVTFVITAGP--LATLT 2818

Query: 552 FNVSPSDIVADGSMQSILTFVPRNKNEFVSGITDLEFIQSGVPVTISP-VTENADNYTA 610
+ ++ AD S + + ++ + + G +G +T + NA+
Sbjct: 2819 LVTTVNNAFADNSATNTVQATLKDVSGNPVGEVVAFAASNGATITATDGGVSNANGIVL 2878

Query: 611 SVVGNSVGDVDITPQVGGESLDLLQKRITLYPVPKITGINVNGEQFATDKGFPKTTFNKA 670
+ + N V V L + T + K + VNG F D GFP T F A
Sbjct: 2879 ATLTNGTAGVS---TVTATITETLTETDTTFIAMKNLDVTVNGTTFNGDAGFPTTGfVGA 2935

Query: 671 TFQLVMNDDVANNTQYDWTSSYAASAPVDNQGV--NIAKYTYGSTVTVTAKSKKFPSYT 728
TF++ D N+ YDW+SS A V G V N + T T+T++A K S
Sbjct: 2936 TFKVNSGGD---NSLYDWSSAPALVSVSGDGVVTFNAVFPPTGTPTITISATPKGGGS-P 2991

Query: 729 ATYQFKPNLWVFSGTMSLQSSVEASRNCQRTDF-----TALIESARASNGSRSPDGTWLG 783
+Y F+ N W + + + +A +C+ + T + +A +G R+ G LW
Sbjct: 2992 LSYSFRVNQWFINNNGATLNRADAITHCENVGYTMTPTSTQVTNAATWMSGKRAV-GNLWS 3050

Query: 784 EWGSLATYDSAEW-PSGNYWTKKT--STDFVTMDMTTGDIPSTA---ATAYPLCAEP 834
EWG + Y + W P+ +W ++ + + ++TG + T A + +C P
Sbjct: 3051 EWGDFSAYTAPGWVPAEFFWLSNNHDASTALAIGLSTGTTLTMGDFMAITHVMCTRP 3107

Score = 60.1 bits (144), Expect = 2e-07
Identities = 54/203 (26%), Positives = 86/203 (41%), Gaps = 7/203 (3%)

Query: 418 AVNTYTLSATAIDNH--GNSSNPATLTVIVQQPQFVITSE---VTDDGALADGRTPITVK 472
A T+TL+ T++ GN T I + I + +T + ALADG +
Sbjct: 2078 ATLTHTLAGTSVVTARVGNRVQSKDTTFIADRTTATIRASDLTITRNNALADGVATNAAR 2137

Query: 473 FTVTNIDSTPVAEQEGVITTSNGALPSKVTKKTDAGQVISIALTSFTVGVSVVTLDI--Q 530
VT+ + PV T+ NGAL + + TD+ G S T T G+S VT I
Sbjct: 2138 VIVTDANGNPVPSMFVGYTSDNGALLTPTSGMTDSSGTFSTTFHTTAGISKVTAIVTM 2197

Query: 531 GQQATVDVRFVAVLPPDVTNSSFNVSPSDIVADGSMQSILTFVPRNKNEFVSGITDLEFI 590
G T D F S V +D +A+ S ++I+ ++ + V+G+
Sbjct: 2198 GISQTKDAVFIADRSTAHVSELIVVKNDLANNSDRNIVQAHIKDAHGNVVTGMNVNFSA 2257

Query: 591 QSGVPVTISPVTENADNYTASVV 613
V +T + VT N+ Y + +

Sbjct: 2258 TENVTLTANTVTTNSQGYAENTL 2280

Score = 56.6 bits (135), Expect = 2e-06

Identities = 63/250 (25%), Positives = 93/250 (37%), Gaps = 10/250 (4%)

Query: 373 NDAQWVAAGGKLIPVSATDYNVVLPPYKPMAPASRTVGKTGESEAAVNTYTLSATAIDNH 432
N Q V + PV T VV A + +G TG A T T + N

Sbjct: 1155 NSVQAVVSDSDGNPV--TGAADVFFSSANATAQITTVIGTTGADGIATATLTNTVAGTSNV 1212

Query: 433 G---NSSNPATLTVIVQQPQFVITSEVTDDGALADGRTPITVKFTVTNIDSTPVAEQEGV 489
++ N T V IT +GA+ADG V V + + P+ V

Sbjct: 1213 VATIDTVNANIDTAFVAGAVATITLTAPVNGAVADGADTNQVDALVEDANGNPITGAADV 1272

Query: 490 ITTSNGALPSKVTKKTDAGQGVISIALTSFTVGVSVVTLDIQQQATVDVRFVAVLPPDVTN 549
+++NGA T T GV S LT G S V + A +D F +P V


Sbjct: 1273 FSSANGATILSSTMNTGVNGVASTLLTHTVAGTSNVVATVDTVNANIDTAF--VPGAVAT 1330

Query: 550 SSFNVSPSIVADGSMQSILTFVPRNKNEFVSGITDLEFIQSGVPVTISPV--TENADN 607
+ + VADG+ + + V + V+G + F + I+ V T AD

Sbjct: 1331 ITLTPVNGAVADGANSNSVQAVVSDSEGNVAVAGAA-VVFSSANATAQITTVIGTTGADG 1389

Query: 608 YTASVVGNSV 617
+ + N+V

Sbjct: 1390 IATATLTNTV 1399

 ExPASy Home page	Site Map	Search ExPASy	Contact us	Swiss-Prot
Hosted by NCSC US Mirror sites: Australia Bolivia Canada China Korea Switzerland Taiwan				
Search <input type="text" value="Swiss-Prot/TrEMBL"/>		<input type="button" value="Go"/>	<input type="button" value="Clear"/>	

Search in Swiss-Prot and TrEMBL for: invasin

Swiss-Prot Release 44.5 of 13-Sep-2004

TrEMBL Release 27.5 of 13-Sep-2004

- Number of sequences found in [Swiss-Prot](#)₍₁₁₎ and [TrEMBL](#)₍₄₅₎: **56**
- Note that the selected sequences can be saved to a file to be later retrieved; to do so, go to the [bottom](#) of this page.
- For more directed searches, you can use the Sequence Retrieval System [SRS](#).

Search in Swiss-Prot: There are matches to 11 out of 158316 entries

INVA_YEREN (P19196)

Invasin. Yersinia enterocolitica

INVA_YERPS (P11922)

Invasin. - Yersinia pseudotuberculosis

IPAA_SHIFL (P18010)

Invasin ipaA (70 kDa antigen). {GENE: Name=ipaA; OrderedLocusNames=CP0125} - Shigella flexneri

IPAB_SHIDY (Q03945)

Invasin ipaB (62 kDa antigen). {GENE: Name=ipaB} - Shigella dysenteriae

IPAB_SHIFL (P18011)

Invasin ipaB (62 kDa antigen). {GENE: Name=ipaB; OrderedLocusNames=CP0128} - Shigella flexneri

IPAC_SHIDY (Q03946)

Invasin ipaC (42 kDa antigen). {GENE: Name=ipaC} - Shigella dysenteriae

IPAC_SHIFL (P18012)

Invasin ipaC (42 kDa antigen). {GENE: Name=ipaC; OrderedLocusNames=CP0127} - Shigella flexneri

IPAD_SHIDY (Q03947)

Invasin ipaD (37 kDa membrane antigen). {GENE: Name=ipaD} - Shigella dysenteriae

IPAD_SHIFL (P18013)

Invasin ipaD (36 kDa membrane antigen). {GENE: Name=ipaD; OrderedLocusNames=CP0126} - Shigella flexneri

YADA_YEREN (P31489)

Invasin precursor (Outer membrane adhesin). {GENE: Name=yadA; Synonyms=yopA, invA, yop1} - Yersinia enterocolitica

YADA_YERPS (P10858)

Invasin precursor (Outer membrane adhesin). {GENE:Name=yadA; Synonyms=yopA, invA, yop1} - Yersinia pseudotuberculosis

Search in TrEMBL: There are matches to 45 out of 1400820 entries

O07390

Invasin 1 {GENE:Name=inv1} - Mycobacterium avium

O07391

Invasin 2 {GENE:Name=inv2} - Mycobacterium avium

O84129

Predicted polysaccharide hydrolase-invasin repeat family {GENE:Name=ydhO;
OrderedLocusNames=CT127} - Chlamydia trachomatis

O84606

Invasin repeat family phosphatase {GENE:Name=papQ; OrderedLocusNames=CT601} -
Chlamydia trachomatis

O84764

Muramidase (Invasin repeat family) {GENE:Name=nlpD; OrderedLocusNames=CT759} -
Chlamydia trachomatis

Q47380

Invasin - Escherichia coli

Q56889

Invasin {GENE:Name=invA} - Yersinia enterocolitica

Q56937

Invasin {GENE:Name=inv} - Yersinia pestis

Q6T8F2

Putative invasin-like protein (Fragment) - Chlamydia suis

Q6T8F6

TetR(C)-invasin fusion protein (Fragment) - Chlamydia suis

Q6T8F7

Invasin-like protein (Fragment) - Chlamydia suis

Q6T8G9

Invasin-like protein (Fragment) - Chlamydia suis

Q74QZ6

Putative invasin {GENE:OrderedLocusNames=YP3306} - Yersinia pestis

Q7A8L6

Putative invasin {GENE:OrderedLocusNames=ECs5290} - Escherichia coli O157:H7

Q7AGX0

Adhesin/invasin-like protein {GENE:OrderedLocusNames=ECs0548} - Escherichia coli
O157:H7

Q7AHB0

Putative invasin {GENE:OrderedLocusNames=ECs0336} - Escherichia coli O157:H7

Q7N592

Similar to putative invasin and adhesin {GENE:OrderedLocusNames=plu2064} - Photorhabdus
luminescens (subsp. laumondii)

Q83T27

Invasin-like protein {GENE:OrderedLocusNames=t1573} - Salmonella typhi

Q8EFM1

Invasin domain protein {GENE:OrderedLocusNames=SO1949} - Shewanella oneidensis

Q8FY77

- Intimin/invasin family protein {GENE:OrderedLocusNames=BR2009} - Brucella suis
Q8KH85
Putative adhesin/invasin - Neisseria meningitidis
- Q8KHF7
Putative adhesin/invasin - Neisseria meningitidis
- Q8KHP5
Putative adhesin/invasin - Neisseria meningitidis
- Q8KI42
Putative adhesin/invasin - Neisseria meningitidis
- Q8KUJ1
Putative adhesin/invasin - Neisseria meningitidis
- Q8KUJ6
Putative adhesin/invasin - Neisseria meningitidis
- Q8KWG5
Putative invasin protein {GENE:Name=agg3B} - Escherichia coli [Plasmid pAA-like]
- Q8RBD9
Soluble lytic murein transglycosylase and related regulatory proteins (Some contain LysM/invasin domains) {GENE:Name=MltE2; OrderedLocusNames=TTE0883} - Thermoanaerobacter tengcongensis
- Q8RBE4
Soluble lytic murein transglycosylase and related regulatory proteins (Some contain LysM/invasin domains) {GENE:Name=MltE; OrderedLocusNames=TTE0876} - Thermoanaerobacter tengcongensis
- Q8RQ63
Putative adhesin/invasin {GENE:Name=Aa32-1-2} - Actinobacillus actinomycetemcomitans (Haemophilus actinomycetemcomitans)
- Q8VM53
Putative invasin {GENE:Name=pagN} - Salmonella enterica IIIb 50:k:z
- Q8XB95
Putative invasin {GENE:OrderedLocusNames=z5932} - Escherichia coli O157:H7
- Q8Z795
Invasin-like protein {GENE:OrderedLocusNames=STY1395} - Salmonella typhi
- Q8Z7G3
Putative invasin {GENE:OrderedLocusNames=STY1284, t1677} - Salmonella typhi
- Q8ZA73
Putative invasin (Hypothetical) {GENE:OrderedLocusNames=YPO3944, y3884} - Yersinia pestis
- Q8ZP33
Putative invasin {GENE:Name=yehP; OrderedLocusNames=STM1768} - Salmonella typhimurium
- Q8ZP78
Homology to invasin C of Yersinia; intimin {GENE:OrderedLocusNames=STM1669} - Salmonella typhimurium
- Q9AM40
Invasin (Fragment) {GENE:Name=inv} - Yersinia pseudotuberculosis
- Q9JSH2
Polysaccharide hydrolase-invasin repeat family {GENE:Name=ydhO} - Chlamydia pneumoniae (Chlamydomydia pneumoniae)
- Q9JXK7
Adhesin/invasin, putative {GENE:OrderedLocusNames=NMB1994} - Neisseria meningitidis (serogroup B)
- Q9KK70

Putative invasin 996A009 (Fragment) - Mycobacterium avium
[Q9L574](#)

Invasin-like SivH {GENE:Name=sivH} - Salmonella typhi
[Q9X4L4](#)

Invasin homolog AafB {GENE:Name=aafB} - Escherichia coli
[Q9Z704](#)

Muramidase (Invasin repeat family) (Hypothetical protein CP0964) (Cell wall hydrolase)
{GENE:Name=nlpD; Synonyms=lytE; OrderedLocusNames=CP0964, CPn0902, CpB0934} -
Chlamydia pneumoniae (Chlamydophila pneumoniae)

[Q9Z8T9](#)
Polysaccharide Hydrolase-Invasin Repeat Family (NLP/P60 family protein) (P60)
{GENE:Name=ydhO; OrderedLocusNames=CP0517, CPn0245, CpB0252} - Chlamydia
pneumoniae (Chlamydophila pneumoniae)

[New Search](#)

in Swiss-Prot/TrEMBL by AC, ID, description,
gene name, organism

**Please do NOT use any boolean operators (and,
or, etc.)**

If you would like to retrieve all the Swiss-Prot/TrEMBL entries contained in this list , you can enter a file name. These entries will then be saved to a file under this name in the directory outgoing of the [ExPASy anonymous ftp server](#), from where you can download it. (Please note that this temporary file will only be kept for 1 week.)

File name:

Format: ☒ Swiss-Prot ☐ Fasta

[Reset](#)

or

[Create file](#)



[ExPASy Home page](#)

[Site Map](#)

[Search ExPASy](#)

[Contact us](#)

[Swiss-Prot](#)

[Hosted by NCSC US](#)

[Mirror sites:](#)

[Australia](#)

[Bolivia](#)


[Canada](#)

[China](#)

[Korea](#)

[Switzerland](#)

[Taiwan](#)

 ExPASy Home page	Site Map	Search ExPASy	Contact us	Swiss-Prot
--	--------------------------	-------------------------------	----------------------------	----------------------------

Search for

NiceProt View of TrEMBL: Q56889

[Printer-friendly view](#)[Request update](#)[Quick BlastP search](#)

[\[Entry info\]](#) [\[Name and origin\]](#) [\[References\]](#) [\[Comments\]](#) [\[Cross-references\]](#) [\[Keywords\]](#)
[\[Features\]](#) [\[Sequence\]](#) [\[Tools\]](#)

Note: most headings are clickable, even if they don't appear as links. They link to the [user manual](#) or [other documents](#).

Entry information

Entry name	Q56889
Primary accession number	Q56889
Secondary accession numbers	None
Entered in TrEMBL in	Release 01, November 1996
Sequence was last modified in	Release 01, November 1996
Annotations were last modified in	Release 26, March 2004
Name and origin of the protein	
Protein name	Invasin
Synonyms	None
Gene name	Name: invA
From	<u>Yersinia enterocolitica</u> [TaxID: 630]
Taxonomy	<u>Bacteria</u> ; <u>Proteobacteria</u> ; <u>Gammaproteobacteria</u> ; <u>Enterobacteriales</u> ; <u>Enterobacteriaceae</u> ; <u>Yersinia</u> .

References

[1] SEQUENCE FROM NUCLEIC ACID.

STRAIN=W1024;

MEDLINE=94195100;PubMed=7511772 [[NCBI](#), [ExPASy](#), [EBI](#), [Israel](#), [Japan](#)]

[Pepe J.C.](#), [Badger J.L.](#), [Miller V.L.](#);

"Growth phase and low pH affect the thermal regulation of the *Yersinia enterocolitica* inv gene."; *Mol. Microbiol.* 11:123-135(1994).

[2] SEQUENCE FROM NUCLEIC ACID.

STRAIN=W1024;

MEDLINE=98048470;PubMed=9387224 [[NCBI](#), [ExPASy](#), [EBI](#), [Israel](#), [Japan](#)]

[Fauconnier A.](#), [Allaoui A.](#), [Campos A.](#), [Van Elsen A.](#), [Cornelis G.R.](#), [Bollen A.](#);

"Flagellar flhA, flhB and flhE genes, organized in an operon, cluster upstream from the inv locus in *Yersinia enterocolitica*."; *Microbiology* 143:3461-3471(1997).

Comments

None

Cross-references

EMBL [Z48169](#); [CAA88188.1](#); -. [[EMBL](#) / [GenBank](#) / [DDBJ](#)] [[CoDingSequence](#)]
 PIR [S54216](#); [S54216](#).
 HSSP [P11922](#); [1CWV](#). [[HSSP ENTRY](#) / [PDB](#)]
[GO:0007155](#); Biological process: cell adhesion (*inferred from electronic annotation*).
 GO [QuickGo](#)
[view](#).
[IPR003344](#); [Big_1](#).
 InterPro [IPR003535](#); [Intimin](#).
[IPR008964](#); [Invasin_intimin](#).
[Graphical view of domain structure](#).
 Pfam [PF02369](#); [Big_1](#); [1](#).
[Pfam graphical view of domain structure](#).
 PRINTS [PR01369](#); [INTIMIN](#).
 ProDom [\[Domain structure / List of seq. sharing at least 1 domain\]](#)
 HOBACGEN [\[Family / Alignment / Tree\]](#)
 ProtoMap [Q56889](#).
 PRESAGE [Q56889](#).
 ModBase [Q56889](#).
 SMR [Q56889](#); [C817F6D7766184E3](#).
 SWISS-2DPAGE [Get region on 2D PAGE](#).
 UniRef View cluster of proteins with at least [50%](#) / [90%](#) identity.

Keywords

None

Features

None

Sequence information

Length: **835** Molecular weight: **91367** CRC64: **C817F6D7766184E3** [This is a checksum on the
 AA Da sequence]

10	20	30	40	50	60
MYSFFNTLTV	TKIISRLILS	IGLIFGIFTY	GFSQQNYFNS	EALENPAEHN	EAFNKIISTG
70	80	90	100	110	120
TSLAVSGNAS	NITRSMVND	ANQEVKHWLN	RFGTTQVNVN	FDKKFSLKES	SLDWLLPWYD
130	140	150	160	170	180
SASYVFFSQL	GIRNKDSRNT	LNIGAGVRTF	QQSWMYGFNT	FYDNDMTGHN	HRIGVGAEAW
190	200	210	220	230	240
TDYLQLSANG	YFRLNGWHQS	RDFVDYNERP	ASGGDIHVKA	YLPALPQLGG	KLKYEQYRGE
250	260	270	280	290	300
RVALFGKDNL	QSNPYAVTTG	LIYTPIPFIT	LGVDQRMGKS	RQHEIQWNLQ	MDYRLGESFR

```

      310           320           330           340           350           360
      |           |           |           |           |           |
SQFSPAVVAG TRLLAESRYN LVERNPNIVL EYQKQNTIKL AFSPAVLSGL PGQVYSVSAQ

      370           380           390           400           410           420
      |           |           |           |           |           |
IQSQSALQRI LWDDAQWIAV GGKLIPVSAT NYNVVLPPYK PMAPASRTVG KTGESEAAVN

      430           440           450           460           470           480
      |           |           |           |           |           |
TYTLSATAID NHGNSSNPAT LTVIVQQPQF VITSEVTDDG ALADGRTPIT VKFTVTNIDG

      490           500           510           520           530           540
      |           |           |           |           |           |
TPVAEQEGVI TTSNGALPSK VTKKTDAGV ISIALTSFTV GGSVVTLDIQ GQQATVDVRF

      550           560           570           580           590           600
      |           |           |           |           |           |
AVLPPDVTNS SFNVSPSDIV ADGSMQSILT FVPRNKNEF VSGITDLEFI QSGVPVTISP

      610           620           630           640           650           660
      |           |           |           |           |           |
VTENADNYTA SVVGNSVGDV DITPQVGGES LDLLQKRITL YPIPKITSIA VNGEQFATDK

      670           680           690           700           710           720
      |           |           |           |           |           |
GFPKTTFNKA TFQLVMNDDV ANNTQYDWTS SYAASAPVDN QGKVNIAYKT YGSTVTVTAK

      730           740           750           760           770           780
      |           |           |           |           |           |
SKKFPSYTAT YQFKPNLWVF SGTMSLQSSI EASRNCQRTD FTALIESARA SNGSRSPDGT

      790           800           810           820           830
      |           |           |           |           |
LWGEWGSLAT YDSAEWPSGN YWTKKTSTDF VTMDMTTGAI PTSAATAYPL CAEPQ

```

Q56889 in [FASTA format](#)

[View entry in original TrEMBL format](#)

[View entry in raw text format \(no links\)](#)

[Request for annotation of this TrEMBL entry](#)

BLAST [BLAST submission on ExPASy/SIB](#)
or at [NCBI \(USA\)](#)



Sequence analysis tools: [ProtParam](#), [ProtScale](#),
[Compute pI/Mw](#), [PeptideMass](#), [PeptideCutter](#),
[Dotlet \(Java\)](#)




[ScanProsite](#), [MotifScan](#)



Search the [SWISS-MODEL Repository](#)

ExPASy Home page	Site Map	Search ExPASy	Contact us	Swiss-Prot
Hosted by NCSC US	Mirror sites:	Australia	Bolivia	Canada
		China	Korea	Switzerland
			Taiwan	

 ExPASy Home page	Site Map	Search ExPASy	Contact us	Swiss-Prot
--	--------------------------	-------------------------------	----------------------------	----------------------------

Search for

NiceProt View of Swiss- Prot: P31489

[Printer-friendly view](#)
[Submit update](#)
[Quick BlastP search](#)

[\[Entry info\]](#) [\[Name and origin\]](#) [\[References\]](#) [\[Comments\]](#) [\[Cross-references\]](#) [\[Keywords\]](#)
[\[Features\]](#) [\[Sequence\]](#) [\[Tools\]](#)

Note: most headings are clickable, even if they don't appear as links. They link to the user manual or other documents.

Entry information

Entry name	YADA_YEREN
Primary accession number	P31489
Secondary accession numbers	None
Entered in Swiss-Prot in	Release 26, July 1993
Sequence was last modified in	Release 26, July 1993
Annotations were last modified in	Release 44, July 2004
Name and origin of the protein	
Protein name	Invasin [Precursor]
Synonym	Outer membrane adhesin
Gene name	Name: yadA Synonyms: yopA, invA, yop1
From	<u>Yersinia enterocolitica</u> [TaxID: 630]
Encoded on	Plasmid pYV.
Taxonomy	<u>Bacteria</u> ; <u>Proteobacteria</u> ; <u>Gammaproteobacteria</u> ; <u>Enterobacteriales</u> ; <u>Enterobacteriaceae</u> ; <u>Yersinia</u> .

References

[1] SEQUENCE FROM NUCLEIC ACID.
STRAIN=6471/76 / Serotype O:3;
 MEDLINE=95020586;PubMed=7934875 [[NCBI](#), [ExPASy](#), [EBI](#), [Israel](#), [Japan](#)]
[Tamm A.](#), [Tarkkanen A.](#), [Korhonen T.K.](#), [Kuusela P.](#), [Toivanen P.](#), [Skurnik M.](#);
 "Hydrophobic domains affect the collagen-binding specificity and surface polymerization as well as the virulence potential of the YadA protein of *Yersinia enterocolitica*."
 Mol. Microbiol. 10:995-1011(1993).

Comments

- **FUNCTION:** Invasin is a protein that allows enteric bacteria to penetrate cultured mammalian cells. The entry of invasins in the cell is mediated by binding several beta-1 chain integrins. This protein is an adhesin forming a fibrillar matrix on the cell surface.
- **SUBCELLULAR LOCATION:** Outer membrane.

Copyright

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)

Cross-references

EMBL X13882; CAA32086.1; -.[\[EMBL / GenBank / DDBJ\]](#) [\[CoDingSequence\]](#)
 PIR S04912; S04912.
[IPR008126](#); Adhesion.
[IPR008640](#); Hep_Hag.
 InterPro [IPR008635](#); HIM.
[IPR005594](#); YadA_C.
[Graphical view of domain structure.](#)
[PF05658](#); Hep_Hag; 4.
 Pfam [PF05662](#); HIM; 1.
[PF03895](#); YadA; 1.
[Pfam graphical view of domain structure.](#)
 PRINTS [PR01756](#); OMADHESIN.
 ProDom [\[Domain structure / List of seq. sharing at least 1 domain\]](#)
 HOBACGEN [\[Family / Alignment / Tree\]](#)
 BLOCKS [P31489](#).
 ProtoNet [P31489](#).
 ProtoMap [P31489](#).
 PRESAGE [P31489](#).
 DIP [P31489](#).
 ModBase [P31489](#).
 SMR [P31489](#); AC12EF68C657DAC0.
 SWISS-2DPAGE [Get region on 2D PAGE](#).
 UniRef [View cluster of proteins with at least 50% / 90% identity.](#)

Keywords

Outer membrane; Plasmid; Signal; Virulence.

Features

[Feature table viewer](#)

Key	From	To	Length	Description
SIGNAL	1	25	25	
CHAIN	26	455	430	Invasin.

Sequence information

Length: **455 AA** [This is the length of the unprocessed precursor]
 Molecular weight: **47136 Da** [This is the MW of the unprocessed precursor]

CRC64: **AC12EF68C657DAC0** [This is a checksum on the sequence]

10	20	30	40	50	60
MTKDFKISVS	AALISALFSS	PYAFADDYDG	IPNLTAVQIS	PNADPALGLE	YPVRPPVPGA
70	80	90	100	110	120

```

      |      |      |      |      |      |
GGLNASAKGI HSAIGATAE AAKGAAVAVG AGSIATGVNS VAIGPLSKAL GDSAVTYGAA
      |      |      |      |      |      |
      130      140      150      160      170      180
STAQKDGVAI GARASTSDTG VAVGFNSKAD AKNSVAIGHS SHVAANHGYG IAIGDRSKTD
      |      |      |      |      |      |
      190      200      210      220      230      240
RENSVSIGHE SLNRQLTHLA AGTKDTPDAVN VAQLKKEIEK TQENTNKRSA ELLANANAYA
      |      |      |      |      |      |
      250      260      270      280      290      300
DNKSSSVLGI ANNYTDSKSA ETLENARKEA FAQSKDVLNM AKAHSNSVAR TTLETAEHHA
      |      |      |      |      |      |
      310      320      330      340      350      360
NSVARTTLET AEEHANKKSA EALASANVYA DSKSSHTLKT ANSYTDVTVS NSTKKAIRES
      |      |      |      |      |      |
      370      380      390      400      410      420
NQYTDHKFRQ LDNRLDKLDT RVDKGLASSA ALNSLFQPYG VGKVNFTAGV GGYRSSQALA
      |      |      |      |      |      |
      430      440      450
IGSGYRVNEN VALKAGVAYA GSSDVMYNAS FNIEW

```

P31489 in [FASTA format](#)

[View entry in original Swiss-Prot format](#)

[View entry in raw text format \(no links\)](#)

[Report form for errors/updates in this Swiss-Prot entry](#)

BLAST [BLAST submission on ExPASy/SIB](#)
or at [NCBI \(USA\)](#)




Sequence analysis tools: [ProtParam](#), [ProtScale](#),
[Compute pI/Mw](#), [PeptideMass](#), [PeptideCutter](#),
[Dotlet \(Java\)](#)




[ScanProsite](#), [MotifScan](#)



Search the [SWISS-MODEL Repository](#)

 ExPASy Home page	Site Map	Search ExPASy	Contact us	Swiss-Prot				
Hosted by NCSC US	Mirror sites:	Australia	Bolivia	Canada	China	Korea	Switzerland	Taiwan

 ExPASy Home page	Site Map	Search ExPASy	Contact us	Swiss-Prot
--	--------------------------	-------------------------------	----------------------------	----------------------------

Search for

NiceProt View of Swiss- Prot: P10858

[Printer-friendly view](#)
[Submit update](#)
[Quick BlastP search](#)

[\[Entry info\]](#)
[\[Name and origin\]](#)
[\[References\]](#)
[\[Comments\]](#)
[\[Cross-references\]](#)
[\[Keywords\]](#)
[\[Features\]](#)
[\[Sequence\]](#)
[\[Tools\]](#)

Note: most headings are clickable, even if they don't appear as links. They link to the user manual or other documents.

Entry information

Entry name	YADA_YERPS
Primary accession number	P10858
Secondary accession numbers	None
Entered in Swiss-Prot in	Release 11, July 1989
Sequence was last modified in	Release 11, July 1989
Annotations were last modified in	Release 44, July 2004
Name and origin of the protein	
Protein name	Invasin [Precursor]
Synonym	Outer membrane adhesin
Gene name	Name: yadA
	Synonyms: yopA, invA, yop1
From	<u>Yersinia pseudotuberculosis</u> [TaxID: 633]
Encoded on	Plasmid pIB1.
Taxonomy	<u>Bacteria</u> ; <u>Proteobacteria</u> ; <u>Gammaproteobacteria</u> ; <u>Enterobacteriales</u> ; <u>Enterobacteriaceae</u> ; <u>Yersinia</u> .

References

- [1] SEQUENCE FROM NUCLEIC ACID.
STRAIN=YPIII;
 DOI=[10.1038/334522a0](#);MEDLINE=88302441;PubMed=3043229 [[NCBI](#), [ExPASy](#), [EBI](#), [Israel](#), [Japan](#)]
Rosqvist R., Skurnik M., Wolf-Watz H.;
 "Increased virulence of *Yersinia pseudotuberculosis* by two independent mutations.";
 Nature 334:522-525(1988).
- [2] SEQUENCE FROM NUCLEIC ACID.
STRAIN=YPIII;
 MEDLINE=89343638;PubMed=2761389 [[NCBI](#), [ExPASy](#), [EBI](#), [Israel](#), [Japan](#)]
Skurnik M., Wolf-Watz H.;
 "Analysis of the yopA gene encoding the Yop1 virulence determinants of *Yersinia* spp.";
 Mol. Microbiol. 3:517-529(1989).

Comments

- **FUNCTION:** Invasin is a protein that allows enteric bacteria to penetrate cultured mammalian cells. The entry of invasin in the cell is mediated by binding several beta-1 chain integrins. This protein is an adhesin forming a fibrillar matrix on the cell surface.
- **SUBCELLULAR LOCATION:** Outer membrane.

Copyright

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)

Cross-references

EMBL	X13883; CAA32088.1; -. [EMBL / GenBank / DDBJ] [CoDingSequence]
PIR	S04534; S04534. IPR008126 ; Adhesion. IPR008640 ; Hep_Hag.
InterPro	IPR008635 ; HIM. IPR005594 ; YadA_C. Graphical view of domain structure.
Pfam	PF05658 ; Hep_Hag; 4. PF05662 ; HIM; 1. PF03895 ; YadA; 1. Pfam graphical view of domain structure.
PRINTS	PR01756 ; OMADHESIN.
ProDom	[Domain structure / List of seq. sharing at least 1 domain]
HOBACGEN	[Family / Alignment / Tree]
BLOCKS	P10858 .
ProtoNet	P10858 .
ProtoMap	P10858 .
PRESAGE	P10858 .
DIP	P10858 .
ModBase	P10858 .
SMR	P10858 ; EE2C55FB12B183D4.
SWISS-2DPAGE	Get region on 2D PAGE .
UniRef	View cluster of proteins with at least <u>50%</u> / <u>90%</u> identity.

Keywords

Outer membrane; Plasmid; Signal; Virulence.

Features



[Feature table viewer](#)

Key	From	To	Length	Description
SIGNAL	<u>1</u>	<u>25</u>	25	
CHAIN	<u>26</u>	<u>434</u>	409	Invasin.

Sequence information

Length: **434 AA** [This is the length of the unprocessed
Molecular weight: **45054 Da** [This is the MW of the

CRC64: **EE2C55FB12B183D4** [This is a checksum on the sequence]


```

precursor]                unprocessed precursor]
      10      20      30      40      50      60
      |       |       |       |       |       |
MTKDFKISVS AALISALFSS PYAFAEPEP GNDGIPRLSA VQISPNVDPK LGVGLYPAKP
      70      80      90     100     110     120
      |       |       |       |       |       |
ILRQENPKLP PRGPQGPEKK RARLAEAIQP QVLGAGGLNA RAKDPYSIAI GATAEAAKPA
     130     140     150     160     170     180
      |       |       |       |       |       |
AVAVGSGSIA TGVNSVAIGP LSKALGDSAV TYGASSTAQK DGVAIGARAS ASDTGVAVGF
     190     200     210     220     230     240
      |       |       |       |       |       |
NSKVDAQNSV AIGHSSHVAA DHGYSIAIGD HSKTDRENSV SIGHESLNRQ LTHLAAGTED
     250     260     270     280     290     300
      |       |       |       |       |       |
TDAVNVAQLK KEMAETLENA RKETLAQSDN VLDAACKHSN SVARTTLETA EEHANKKSAE
     310     320     330     340     350     360
      |       |       |       |       |       |
ALVSAKVYAD SNSSHTLKTA NSYTDVTVSS STKKAISESN QYTDHKFSQL DNRLDKLDR
     370     380     390     400     410     420
      |       |       |       |       |       |
VDKGLASSAA LNSLFQPYGV GKVNF TAGVG GYRSSQALAI GSGYRVNESV ALKAGVAYAG
     430
      |
SSNVMYNASF NIEW

```

P10858 in [FASTA format](#)

[View entry in original Swiss-Prot format](#)

[View entry in raw text format \(no links\)](#)

[Report form for errors/updates in this Swiss-Prot entry](#)

BLAST [BLAST submission on ExPASy/SIB](#)
or at [NCBI \(USA\)](#)




Sequence analysis tools: [ProtParam](#), [ProtScale](#),
[Compute pI/Mw](#), [PeptideMass](#), [PeptideCutter](#),
[Dotlet \(Java\)](#)



[ScanProsite](#), [MotifScan](#)



Search the [SWISS-MODEL Repository](#)

 ExPASy Home page	Site Map	Search ExPASy	Contact us	Swiss-Prot				
Hosted by NCSC US	Mirror sites:	Australia	Bolivia	Canada	China	Korèa	Switzerland	Taiwan

 ExPASy Home page	Site Map	Search ExPASy	Contact us	Proteomics tools	Swiss-Prot
--	--------------------------	-------------------------------	----------------------------	----------------------------------	----------------------------

Search for

=====

Welcome to the SIB BLAST Network Service

If results of this search are reported or published, please mention that the computation was performed at the SIB using the BLAST network service. The SIB BLAST network service uses a server developed at SIB and the NCBI BLAST 2 software.

In case of problems, please read the [online BLAST help](#).
If your question is not covered, please contact [<helpdesk@expasy.org>](mailto:helpdesk@expasy.org).

NCBI BLAST program reference [PMID:9254694]:

Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389-3402 (1997).

=====

Query length: 682 AA (of which 2% low-complexity regions filtered out)

Date run: 2004-09-14 18:42:59 UTC+0100 on sib-gm1.unil.ch

Program: NCBI BLASTP 1.5.4-Paracel [2003-06-05]

Database: EXPASY/UniProt

1,568,464 sequences; 501,960,298 total letters

[Taxonomic view](#)

[NiceBlast view](#)

[Printable view](#)

List of potentially matching sequences

Send selected sequences to

☐ Include query sequence

Db AC	Description	Score	E-value
<input type="checkbox"/> tr Q51112	Similar to Shigella flexneri VirG protein (Fragment) [...	1367	0.0
<input type="checkbox"/> tr Q9JTL4	Putative virulence associated protein [NMA1725] [Neiss...	1043	0.0
<input type="checkbox"/> tr Q6Q509	Lav (Fragment) [Haemophilus influenzae]	962	0.0
<input type="checkbox"/> tr Q93N95	Lav (Fragment) [lav] [Haemophilus influenzae]	954	0.0
<input type="checkbox"/> tr Q93N94	Las (Fragment) [las] [Haemophilus influenzae biotype a...	663	0.0
<input type="checkbox"/> tr Q9JPL5	Autotransporter A [autA] [Neisseria meningitidis]	389	e-107
<input type="checkbox"/> tr Q9ZF58	Virulence-associated protein VapA [vapA] [Neisseria me...	388	e-106
<input type="checkbox"/> tr Q9JSR2	Putative virulence associated protein [vapA] [Neisseri...	387	e-106
<input type="checkbox"/> tr Q6J5G0	Putative virulence-associated protein (Fragment) [Haem...	219	2e-55
<input type="checkbox"/> tr Q8X6C1	Putative beta-barrel outer membrane protein [z0402] [E...	142	2e-32

<input type="checkbox"/>	tr Q7AH89	AidA-I adhesin-like protein [ECs0362] [Escherichia col...	142	2e-32
<input type="checkbox"/>	tr Q6KCW6	YapH homolog [yapH] [Escherichia coli]	139	1e-31
<input type="checkbox"/>	tr Q8FFF9	Hypothetical protein ydeU [ydeU] [Escherichia coli O6]	137	9e-31
<input type="checkbox"/>	tr Q7BCK4	VirG [virG] [Shigella flexneri]	134	4e-30
<input type="checkbox"/>	tr Q99Q93	IcsA (VirG), outer membrane protein exposed to the bact...	134	4e-30
<input type="checkbox"/>	tr Q52298	Virulence-associated VirG [Plasmid pMYSH6000]	134	4e-30
<input type="checkbox"/>	tr Q9JMS3	YchA protein [ychA] [Escherichia coli]	134	6e-30
<input type="checkbox"/>	tr Q6WE27	IcsA [icsA] [Shigella flexneri]	134	7e-30
<input type="checkbox"/>	tr Q8ZHA1	Putative autotransporter protein [yapH] [Yersinia pestis]	130	1e-28
<input type="checkbox"/>	tr Q8CZU2	Putative autotransporter adhesin [yapH] [Yersinia pestis]	130	1e-28
<input type="checkbox"/>	tr Q9F289	YapD protein [yapD] [Yersinia pestis]	130	1e-28
<input type="checkbox"/>	tr Q9F285	YapH protein [yapH] [Yersinia pestis]	130	1e-28
<input type="checkbox"/>	tr Q74QP7	Putative autotransporter protein [yapH] [Yersinia pestis]	130	1e-28
<input type="checkbox"/>	tr Q9F286	YapG protein (Putative ATP-binding transport component...	129	2e-28
<input type="checkbox"/>	tr Q74RV9	Putative autotransporter protein [YP2907] [Yersinia pe...	129	2e-28
<input type="checkbox"/>	sp P77286	YDEU_ECOLI Hypothetical protein ydeU [ydeU] [Escherich...	129	2e-28
<input type="checkbox"/>	tr Q8XAY3	Putative ATP-binding component of a transport system a...	129	2e-28
<input type="checkbox"/>	tr Q7AE01	Putative ATP-binding component of a transport system a...	129	2e-28
<input type="checkbox"/>	tr Q83LF4	Hypothetical protein SF1157 [SF1157] [Shigella flexneri]	128	4e-28
<input type="checkbox"/>	tr Q7UCU4	Hypothetical protein [S1242] [Shigella flexneri]	128	4e-28
<input type="checkbox"/>	tr Q9CKA8	Hypothetical protein PM1717 [PM1717] [Pasteurella mult...	126	1e-27
<input type="checkbox"/>	tr P75997	Putative part of putative ATP-binding component of a t...	118	3e-25
<input type="checkbox"/>	sp Q03155	AIDA_ECOLI Adhesin aidA-I precursor [aidA-I] [Escheric...	118	4e-25
<input type="checkbox"/>	tr Q7CPG9	Putative autotransported protein [misL] [Salmonella ty...	110	7e-23
<input type="checkbox"/>	tr Q9Z625	MisL [misL] [Salmonella typhimurium]	110	7e-23
<input type="checkbox"/>	sp P45508	YFAL_ECOLI Hypothetical protein yfaL precursor [yfaL] ...	96	2e-18
<input type="checkbox"/>	tr Q8ZN57	Similar to the C-terminal region of AIDA [shdA] [Salmo...	96	3e-18
<input type="checkbox"/>	tr Q8CVV7	Hypothetical protein yfaL [yfaL] [Escherichia coli O6]	96	3e-18
<input type="checkbox"/>	tr Q9XCJ4	ShdA [shdA] [Salmonella typhimurium]	96	3e-18
<input type="checkbox"/>	tr Q8XE28	Putative ATP-binding component of a transport system [...	93	1e-17
<input type="checkbox"/>	tr Q7AC44	Putative ATP-binding component of a transport system [...	93	1e-17
<input type="checkbox"/>	tr Q9FCW0	Hypothetical protein [Kluyvera ascorbata]	74	9e-12
<input type="checkbox"/>	tr Q6IU25	YdeU (Fragment) [Escherichia coli B]	70	1e-10
<input type="checkbox"/>	tr Q88LP9	Outer membrane autotransporter [PP1880] [Pseudomonas p...	66	2e-09
<input type="checkbox"/>	tr Q884S7	Autotransporter, putative [PSPT02011] [Pseudomonas syr...	66	2e-09
<input type="checkbox"/>	tr Q7VYJ2	Autotransporter [BP1344] [Bordetella pertussis]	65	3e-09
<input type="checkbox"/>	tr Q7WIM1	Putative autotransporter [BB2830] [Bordetella bronchis...	64	7e-09
<input type="checkbox"/>	tr Q883X2	Autotransporter, putative [PSPT02225] [Pseudomonas syr...	64	9e-09
<input type="checkbox"/>	tr Q7WIB0	Putative autotransporter [BB2941] [Bordetella bronchis...	55	3e-06
<input type="checkbox"/>	tr Q7W6E1	Putative autotransporter [BPP2975] [Bordetella parapap...	55	3e-06
<input type="checkbox"/>	tr Q7WQ80	Autotransporter [BB0452] [Bordetella bronchiseptica (A...	54	1e-05
<input type="checkbox"/>	tr Q7WHU5	Autotransporter [BB3111] [Bordetella bronchiseptica (A...	54	1e-05
<input type="checkbox"/>	tr Q7WC76	Autotransporter [BPP0452] [Bordetella parapertussis]	54	1e-05
<input type="checkbox"/>	tr Q8VV95	Vag8 protein (Autotransporter) [vag8-2] [Bordetella pe...	52	4e-05
<input type="checkbox"/>	tr Q66044	Vag8 [vag-8] [Bordetella pertussis]	52	4e-05

<input type="checkbox"/>	tr	Q7WQ83	Autotransporter [BB0450] [Bordetella bronchiseptica (A...	52	5e-05
<input type="checkbox"/>	tr	Q7WC79	Autotransporter [BPP0449] [Bordetella parapertussis]	52	5e-05
<input type="checkbox"/>	tr	Q7W9Y1	Autotransporter [BPP1618] [Bordetella parapertussis]	52	5e-05
<input type="checkbox"/>	tr	Q9F290	YapC protein (Putative autotransporter) (Putative autot...	51	6e-05
<input type="checkbox"/>	tr	Q88ID0	Outer membrane autotransporter [PP3069] [Pseudomonas p...	51	8e-05
<input type="checkbox"/>	sp	P76017	YCGV_ECOLI Hypothetical protein ycgV [ycgV] [Escherich...	50	1e-04
<input type="checkbox"/>	tr	Q881W9	Autotransporter, putative [PSPT02763] [Pseudomonas syr...	50	2e-04
<input type="checkbox"/>	tr	Q7WL85	Autotransporter [vag8] [Bordetella bronchiseptica (Alc...	50	2e-04
<input type="checkbox"/>	tr	Q7W7U5	Autotransporter [vag8] [Bordetella parapertussis]	50	2e-04
<input type="checkbox"/>	tr	Q7WHU6	Autotransporter [BB3110] [Bordetella bronchiseptica (A...	49	2e-04
<input type="checkbox"/>	tr	Q86135	Tracheal colonization factor precursor [tcfA2] [Bordet...	49	2e-04
<input type="checkbox"/>	tr	Q8VV97	Tcf protein [tcfA3] [Bordetella pertussis]	49	2e-04
<input type="checkbox"/>	tr	Q9F4B3	BapB protein [bapB] [Bordetella pertussis]	49	2e-04
<input type="checkbox"/>	tr	Q8GB90	Tracheal colonization factor [tcfA] [Bordetella pertus...	49	2e-04
<input type="checkbox"/>	tr	Q45343	Tracheal colonization factor [tcfA] [Bordetella pertus...	49	2e-04
<input type="checkbox"/>	tr	Q6U948	Tracheal colonization factor protein [tcfA] [Bordetell...	49	2e-04
<input type="checkbox"/>	tr	Q7WKS2	Putative autotransporter [bapC] [Bordetella bronchisep...	48	5e-04
<input type="checkbox"/>	tr	Q83RP3	Putative adhesion and penetration protein [SF1205] [Sh...	47	0.001
<input type="checkbox"/>	tr	Q7WK68	Autotransporter [BB2270] [Bordetella bronchiseptica (A...	47	0.001
<input type="checkbox"/>	tr	Q7VVD6	Autotransporter (Pseudogene) [bapC] [Bordetella pertus...	47	0.001
<input type="checkbox"/>	tr	Q7C1Y6	Putative adhesion and penetration protein [S1289] [Shi...	47	0.001
<input type="checkbox"/>	tr	Q86044	Putative autotransporter (BapC protein) [bapC] [Bordet...	47	0.001
<input type="checkbox"/>	tr	Q7W8V8	Autotransporter [BPP2022] [Bordetella parapertussis]	47	0.001
<input type="checkbox"/>	tr	Q7W1F5	Autotransporter [BPP0735] [Bordetella parapertussis]	46	0.002
<input type="checkbox"/>	tr	Q7WP64	Autotransporter [BB0821] [Bordetella bronchiseptica (A...	45	0.003
<input type="checkbox"/>	tr	Q6G430	Hypothetical protein [BH05490] [Bartonella henselae (R...	44	0.008
<input type="checkbox"/>	tr	Q7W7D4	Putative autotransporter [bapC] [Bordetella parapertus...	44	0.008
<input type="checkbox"/>	tr	Q9JMS5	YcbB protein [ycbB] [Escherichia coli]	44	0.013
<input type="checkbox"/>	tr	Q9XD84	TibA [tibA] [Escherichia coli]	43	0.017
<input type="checkbox"/>	tr	Q6G428	Hypothetical protein [BH05510] [Bartonella henselae (R...	42	0.029
<input type="checkbox"/>	tr	Q83JR2	Exported serine protease SigA [sigA] [Shigella flexneri]	42	0.029
<input type="checkbox"/>	tr	Q9L8L1	Exported serine protease SigA [sigA] [Shigella flexner...	42	0.029
<input type="checkbox"/>	tr	Q7C013	Serine protease [sigA] [Shigella flexneri]	42	0.029
<input type="checkbox"/>	tr	Q8UJW1	Autotransporter protein [yapE] [Agrobacterium tumefaci...	40	0.14
<input type="checkbox"/>	tr	Q7D396	AGR_pAT_528p [AGR_pAT_528] [Agrobacterium tumefaciens ...	40	0.14
<input type="checkbox"/>	tr	Q7WK90	Autotransporter [phg] [Bordetella bronchiseptica (Alca...	40	0.19
<input type="checkbox"/>	tr	Q7W8X9	Autotransporter [phg] [Bordetella parapertussis]	40	0.19
<input type="checkbox"/>	tr	Q84GK0	Secreted autotransporter protein EatA [eatA] [Escheric...	40	0.19
<input type="checkbox"/>	tr	Q8UJX1	Autotransporter protein [bapA] [Agrobacterium tumefaci...	39	0.25
<input type="checkbox"/>	tr	Q7D3A2	AGR_pAT_511p [AGR_pAT_511] [Agrobacterium tumefaciens ...	39	0.25
<input type="checkbox"/>	tr	Q6MTD5	Hypothetical transmembrane protein [MSC_0473] [Mycopla...	39	0.32
<input type="checkbox"/>	sp	P45386	IGA4_HAEIN Immunoglobulin A1 protease precursor (EC 3....	39	0.42
<input type="checkbox"/>	tr	Q9Z5R3	Phg protein (Autotransporter) [phg] [Bordetella pertus...	38	0.55
<input type="checkbox"/>	tr	Q6KD18	Antigen 43 precursor Sap [sap] [Escherichia coli]	38	0.71
<input type="checkbox"/>	tr	Q8FDW4	Accreted auto transpoter toxin [sat] [Escherichia coli...	37	1.2

Graphical overview of the alignments

[Click here](#)

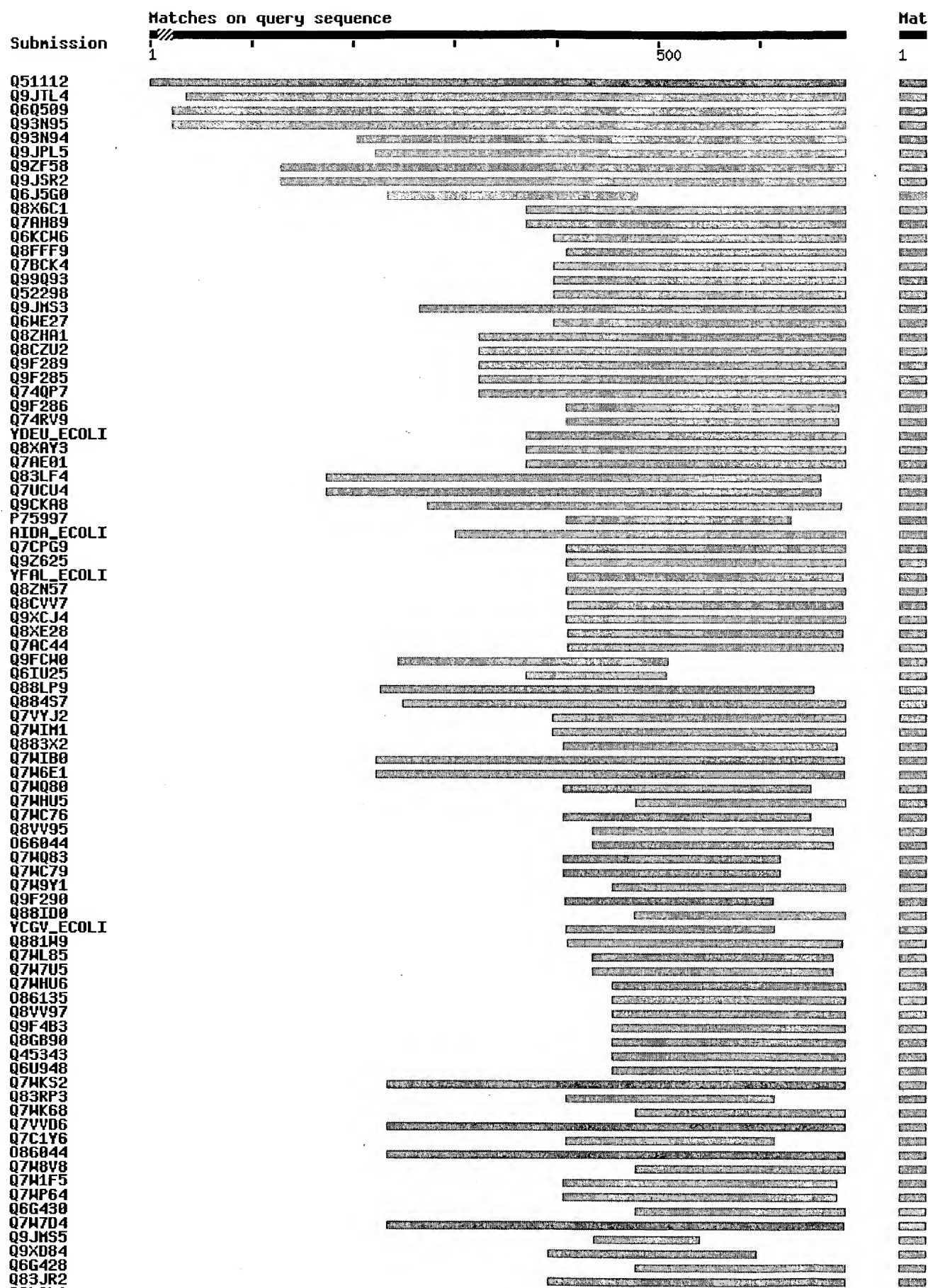
to resubmit your query after masking regions matching PROSITE profiles or Pfam HMMs

([?](#) Help) (use ScanProsite for more details about PROSITE matches)

Profile hits

Pfam hits

Autotransporter



Alignments

tr Q51112 Similar to Shigella flexneri VirG protein (Fragment) 682
[Neisseria AA
meningitidis] align

Score = 1367 bits (3538), Expect = 0.0
Identities = 669/682 (98%), Positives = 669/682 (98%)

Query: 1 QRKIYYETRXXXXXXXXXXXXXIVSLFFSILYTSPLLAVDYVYDKTKLTNDEITRLKKLR 60
QRKIYYETR IVSLFFSILYTSPLLAVDYVYDKTKLTNDEITRLKKLR
Sbjct: 1 QRKIYYETR SKQASKQKFKKSFIVSLFFSILYTSPLLAVDYVYDKTKLTNDEITRLKKLR 60

Query: 61 DKTSEYWKKETYLTEDNPKVPPFPALYPRTYQFENINNSKKISFYDQEYTEGYLVGFAR 120
DKTSEYWKKETYLTEDNPKVPPFPALYPRTYQFENINNSKKISFYDQEYTEGYLVGFAR
Sbjct: 61 DKTSEYWKKETYLTEDNPKVPPFPALYPRTYQFENINNSKKISFYDQEYTEGYLVGFAR 120

Query: 121 GLGVAKRNGDTEEQIRKYFKECFNSNTKIRDYSTCQAEKFGSHPLIVKSHIFSLGPKIKN 180
GLGVAKRNGDTEEQIRKYFKECFNSNTKIRDYSTCQAEKFGSHPLIVKSHIFSLGPKIKN
Sbjct: 121 GLGVAKRNGDTEEQIRKYFKECFNSNTKIRDYSTCQAEKFGSHPLIVKSHIFSLGPKIKN 180

Query: 181 SHINSEILSVGNYTEWANQVIHHIENYVSFAAHLYSGLDPFHYIEVTDNSHVIGQTISLD 240
SHINSEILSVGNYTEWANQVIHHIENYVSFAAHLYSGLDPFHYIEVTDNSHVIGQTISLD
Sbjct: 181 SHINSEILSVGNYTEWANQVIHHIENYVSFAAHLYSGLDPFHYIEVTDNSHVIGQTISLD 240

Query: 241 EFRLNSLWEPRWDSNVGKLTNNADIRFNTKSESLLVKEDYAGGARFRFAYDPKEAKNT 300
EFRLNSLWEPRWDSNVGKLTNNADIRFNTKSESLLVKEDYAGGARFRFAYDPKEAKNT
Sbjct: 241 EFRLNSLWEPRWDSNVGKLTNNADIRFNTKSESLLVKEDYAGGARFRFAYDPKEAKNT 300

Query: 301 ALIFEKNVTGTSDIIFENPIDDLKSLDGHQIIKVNGTADKHAFLRSGKHQKGIYTLSQLQ 360
ALIFEKNVTGTSDIIFENPIDDLKSLDGHQIIKVNGTADKHAFLRSGKHQKGIYTLSQLQ
Sbjct: 301 ALIFEKNVTGTSDIIFENPIDDLKSLDGHQIIKVNGTADKHAFLRSGKHQKGIYTLSQLQ 360

Query: 361 RPEGFLPKCKNAMIWRFMHNRLKPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSN 420
RPEGFLPKCKNAMIWRFMHNRLKPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSN
Sbjct: 361 RPEGFLPKCKNAMIWRFMHNRLKPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSN 420

Query: 421 QWVQGKTAPVEGYRKGVLGGVFTWQNESNQLSIGLMGGQAEQRSTFHNPDNDLTTGN 480
QWVQGKTAPVEGYRKGVLGGVFTWQNESNQLSIGLMGGQAEQRSTFHNPDNDLTTGN
Sbjct: 421 QWVQGKTAPVEGYRKGVLGGVFTWQNESNQLSIGLMGGQAEQRSTFHNPDNDLTTGN 480

Query: 481 VKGFGAGVYATWHQLQDKQTGAYADSWMQYQFRHRINTEDGTERFTSKGITASIEAGYN 540
VKGFGAGVYATWHQLQDKQTGAYADSWMQYQFRHRINTEDGTERFTSKGITASIEAGYN
Sbjct: 481 VKGFGAGVYATWHQLQDKQTGAYADSWMQYQFRHRINTEDGTERFTSKGITASIEAGYN 540

Query: 541 ALLAEHFTKKGNLSRVYLQPQAQLTYLGVNGKFSSENHVNLLGSRQLQTRVGVQAKAQ 600
ALLAEHFTKKGNLSRVYLQPQAQLTYLGVNGKFSSENHVNLLGSRQLQTRVGVQAKAQ
Sbjct: 541 ALLAEHFTKKGNLSRVYLQPQAQLTYLGVNGKFSSENHVNLLGSRQLQTRVGVQAKAQ 600

Query: 601 FSLYKNIAIEPFAAVNALYHNKPFVEMDGERRVINNKTAIESQLGVAVKIKSHLTQAT 660
FSLYKNIAIEPFAAVNALYHNKPFVEMDGERRVINNKTAIESQLGVAVKIKSHLTQAT
Sbjct: 601 FSLYKNIAIEPFAAVNALYHNKPFVEMDGERRVINNKTAIESQLGVAVKIKSHLTQAT 660

Query: 661 FNRQTGKHHQAKQGALNLQWTF 682
FNRQTGKHHQAKQGALNLQWTF
Sbjct: 661 FNRQTGKHHQAKQGALNLQWTF 682

tr Q9JTL4 Putative virulence associated protein [NMA1725] [Neisseria 642 AA
meningitidis (serogroup A)]

align

Score = 1043 bits (2698), Expect = 0.0

Identities = 529/650 (81%), Positives = 568/650 (87%), Gaps = 12/650 (1%)

Query: 37 LAVDYVYDKTKLTNDEITRLKKLRDKTSEYWKKETYLITEDNPKVPPFPALYPRTY--QF 94
+AVDYVYDKTKLT+DEITRLKKLRD+ SEYWK+ETY I +N P PAL+P+ F
Sbjct: 1 MAVDYVYDKTKLTDDEITRLKKLRDRNSEYWKETIYHIKSNNRVYPNIPALFPKHFPDFP 60

Query: 95 ENINNSKKISFYDQEYTEGYLVGFARGLGVAKRNGDTEEQIRKYFKECFNSNTKIRDYST 154
ENINNSK+ISFYD+EYTE YLVGFA+GLGVAKRNG+TE+ IR+YFKEC N+ D T
Sbjct: 61 ENINNSKRISFYDKEYTEDYLVGFAQGLGVAKRNGETEKPIRQYFKECLNTGKYSDD--T 118

Query: 155 CQAEKFGSHPLIVKSHIFSLGPKIKNSHINSEILSVGNYTEWANQVIHHIENYVSFAAHL 214
C++++ S P V+S IF+L KIKNSHINSEILSVGNY EW ++ + S HL
Sbjct: 119 CKSQQ--SIPT-VRSDIFALNTKIKNSHINSEILSVGNYIEWLRPTLNQLS---SSQEH 172

Query: 215 YSGLDPFPHYIEVTDNSHVIGQTISLDEFRLNSLWEPWDSNVGKLKTTNADIRFNTKSE 274
YS +DPFPHYIEVTDNSHVIGQTISLDEFRLNSLWEPWDS+VG+LKTNTNADIRFNTKSE
Sbjct: 173 YSDVDPFPHYIEVTDNSHVIGQTISLDEFRLNSLWEPWDSVVGELKTTNADIRFNTKSE 232

Query: 275 SLLVKEDYAGGARFRFAYDPKEA--KNTALIFEKNVTGTSDIIFENPIDDLKSLDGHQII 332
SLLVKEDYAGGARFRFAY K+ + L FEKN+TGTSDIIFENPIDDLKSLDGHQII
Sbjct: 233 SLLVKEDYAGGARFRFAYGLKDKVPETPVLTFEKNITGTSDIIFENPIDDLKSLDGHQII 292

Query: 333 KVNGTADKHAFRLSGKHQKGIYTLTLQQRPEGFLPKCKNAMIWRFMHNRLKPPIPYCALR 392
KVNGTADKHAFRLSGKHQKGIYTLTLQQRPEGF K + + + ALR
Sbjct: 293 KVNGTADKHAFRLSGKHQKGIYTLTLQQRPEGFFTKVQERDDISIYAQQAQAANTLFLALR 352

Query: 393 LNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGKTAPVEGYRKGVLGGEVFTWQNESNQ 452
LN+KNSDIFDRTLPRKGLWLRVIDGHSNQWVQGKTAPVE RKGVLGGEVFTWQNESNQ
Sbjct: 353 LNDKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGKTAPVESNRKGVLGGEVFTWQNESNQ 412

Query: 453 LSIGLMGGQAEQRSTFHNPDNDLTGTVKGFAGVYATWHQLQDKQTGAYADSWMQYQR 512
LS+GLM GQAEQRSTF NPDTDNDLTGTVKGFAG+YATWHQLQDKQTGAYADSW+QYQR
Sbjct: 413 LSVGLMSGQAEQRSTFRNPDTDNDLTGTVKGFAGIYATWHQLQDKQTGAYADSWVQYQR 472

Query: 513 FRHRINTEDGTERFTSKGITASIEAGYNALLAEHFTKKGNLSRVYLQPQAQLTYLGVNGK 572
FRHRINTED TERFTSKGITASIEAGYNALLAEHFTKKGN +R YLQPQAQLTYLGVNGK
Sbjct: 473 FRHRINTEDATERFTSKGITASIEAGYNALLAEHFTKKGNRVRYLQPQAQLTYLGVNGK 532

Query: 573 FSDSENAHVNLLGSRQLQTRVGVQAKAQFSLYKNIAIEPFAAVNALYHNKPFVEMDGER 632
FSDSENAHVNLLGSRQLQ+RVGVQAKAQFSLYKNIAIEPFAAVNALYHNKPFVEMDGER
Sbjct: 533 FSDSENAHVNLLGSRQLQSRVGVQAKAQFSLYKNIAIEPFAAVNALYHNKPFVEMDGER 592

Query: 633 RVINNKAIAESQLGVAVKIKSHLTLQATFNROTGKHHQAKQGALNLQWTF 682
R+INNKAIAESQLGVAVKIKSHLTLQATFNROTGKHHQAKQGALNLQWTF
Sbjct: 593 RMINNKAIAESQLGVAVKIKSHLTLQATFNROTGKHHQAKQGALNLQWTF 642

tr Q6Q509 Lav (Fragment) [Haemophilus influenzae]

695 AA

align

Score = 962 bits (2486), Expect = 0.0

Identities = 494/665 (74%), Positives = 547/665 (81%), Gaps = 13/665 (1%)

```

Query: 23  IVSLFFSILYTSPLLA-VYVYDKTKLTNDEITRLKLRDKTSEYWKKETYLITEDNPKV 81
      I+SLFFSILYTSPLLA DY YD +KLTN++I RLKLRD+ SEYWKKETYL+
Sbjct: 39  IISLFFSILYTSPLLAVDYTYDNSKLTNEQIERLKKLRDRNSEYWKKETYLKSPSNF 98

Query: 82  PPPALYPR--TYQFENINNSKISFYDQYETEGYLVGFARGLGVAKRNGDTEEQIRKYF 139
      P P L+P+ + FENI+NSK ISFYDQ+YTE YLVGFA+G GVAKRNGDTEE +RKYF
Sbjct: 99  PDIPVLFPKDSSVPFENIDNSKAISFYDQKYTEDYLVGFAQGFGVAKRNGDTEEPVRKYF 158

Query: 140  KECFNSNTKIRDYSTCQAEKFGSHPLIVKSHIFSLGPKIKNSHINSEILSVGNYTEWANQ 199
      KEC N+ D C+A S + +KS IF+ P IKNSHINSEIL+VGNYTE
Sbjct: 159  KECLNTGNYNND--NCKANPLASS-VSIKSDIFTPRPTIKNSHINSEILAVGNYTEL--- 212

Query: 200  VIHHIENYVSFAAHLYSGLDPFHYIEVTDNSHVIGQTISLDEFRLNSLWEPRWDSNVGK 259
      ++ + S+A HLYS DP + V DNSHVIGQTI L +L NSLWEPRW+SN+
Sbjct: 213  MLAAQPSASSWAHLYS--DPGLSLTVQDNSHVIGQTIDLGFLQLTNSLWEPRWNSNIDY 270

Query: 260  LKTTNADIRFNTKSESLLVKEDYAGGARFRFAYDPKEA--KNTALIFEKNVTGTSDIIFE 317
      L T NA+IRFNTK+ESLLV+ YAGGARFRFAYD K+ + L FE+N+TGTSDIIFE
Sbjct: 271  LVTENAEIRFNTKNESLLVEGYAYAGGARFRFAYDLKDKAPETPVLTFEQNITGTSDIIFE 330

Query: 318  NPIDDLKSLDGHQIIKVNGTADKHAFRLSGKHQKGIYTLSSLQQRPEGFLPKCKNAMIWRF 377
      NPIDDLKSLDGHQI+KVNGTADK+AFRLSGK++KGIYTLSSLQQR EGF K +
Sbjct: 331  NPIDDLKSLDGHQIVKVNGTADKNAFRLSGKYRKGITYTLSSLQQRLEGFFTKLQERDDIAI 390

Query: 378  MHNRLKPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGKTAPVEGYRKGV 437
      + + ALRLN+KNSDIFDRTLPRKGLWLRVIDGHS+QWVQGKTAP+EGYRKGV
Sbjct: 391  YAQQAQAANTLFALRLNDKNSDIFDRTLPRKGLWLRVIDGHSSQWVQGKTAPLEGYRKGV 450

Query: 438  QLGGEVFTWQNESNQLSIGLMGGQAEQRSTFHNPDNDLTTGNVKGFGAGVYATWHQLQD 497
      QLGGEVFTWQNESNQ S+GLMGGQAEQRSTF NPDTDNLTTGN+KG GAGVYATWHQLQD
Sbjct: 451  QLGGEVFTWQNESNQFSVGLMGGQAEQRSTFRNPDTDNLTTGNMKGLGAGVYATWHQLQD 510

Query: 498  KQTGAYADSWMQYQFRHRINTEDGTERFTSKGITASIEAGYNALLAEHFTKKGNSLRVY 557
      KQTGAY DSW QYQFRHRINTEDGTERFTSKGITASIEAGYNALLAEH T KG +R Y
Sbjct: 511  KQTGAYVDSWAQYQFRHRINTEDGTERFTSKGITASIEAGYNALLAEHLTGKGTQIRFY 570

Query: 558  LQPQAQLTYLGVNGKFSSENHVNLLGSRQLQTRVGVAQAQFSLYKNIAIEPFPAAVNA 617
      LQPQAQLTYLGVNGKFSSEN+ VNLLGSRQLQ+RVGVQAQAQF L KNI I+PFAAVN
Sbjct: 571  LQPQAQLTYLGVNGKFSSENSQVNLLGSRQLQSRVGVAQAQFLLNKNIVIQPFAAVNT 630

Query: 618  LYHNKPFGEVMDGERRVINNKTAIESQLGVAVKIKSHLTLQATFNRTGKHHQAKQGALN 677
      LYH+KPFGEV+DGERRVINNKTAIESQ G+AVKIKSHLTLQATFNRTGKHH AKQGALN
Sbjct: 631  LYHSPFGEVIDGERRVINNKTAIESQFGIAVKIKSHLTLQATFNRTGKHHHAKQGALN 690

Query: 678  LQWTF 682
      LQWTF
Sbjct: 691  LQWTF 695

```

tr Q93N95 Lav (Fragment) [lav] [Haemophilus influenzae] 692 AA
align

Score = 954 bits (2467), Expect = 0.0

Identities = 487/664 (73%), Positives = 543/664 (81%), Gaps = 13/664 (1%)

Query: 23 IVSLFFSILYTSPLLAVDYVYDKTKLTND EITRLKKLRDKTSEYWKKETYLITEDNPKVP 82
I+SL S LY+SPLLAVDYVYDKTKLT+DEITRLKKLRD+ SEYWK+ETY I P
Sbjct: 38 ILSLLVSALYSSPLLAVDYVYDKTKLTDD EITRLKKLRDRDSEYWKEETYFIKSTPQSP 97

Query: 83 PFPALYPTY--QFENINNSKISFYDQEYTEGYLVGFARGLGVAKRNGDTEEQIRKYFK 140
P L+P+ FENINNSK+ISFYD+E+TE YLVGFA+G GVAKRNG+TEE +R+YFK
Sbjct: 98 NIPRLFPKNSFDSFENINNSKEISFYDKEFTEDYLVGFAQGFGVAKRNGETEETVRQYFK 157

Query: 141 ECFNSNTKIRDYSTCQAEKFGSHPLIVKSHIFSLGPKIKNSHINSEILSVGNYTEWANQV 200
EC N+ D TC++ + + +KS IF+L +KNSHINSEIL+VGNYT+
Sbjct: 158 ECLNTGKHSND-PTCKS--YSTDAYNIKSDIFALNTIVKNSHINSEILAVGNYTKMLSA 214

Query: 201 IHHIENYVSFAAHLYSGLDPFHYIEVTDNSHVIGQTI SLDEFRLNSLWEPRWDSNVGKL 260
H ++A HLYS + + V DNSHVIGQTI L L NSLWEPRW+SN+ L
Sbjct: 215 QHSS----TWAEHLYSNAELS--LTVDNSHVIGQTI DLGALILTNSLWEPRWNSNIDYL 268

Query: 261 KTTNADIRFNTKSESLLVKEDYAGGARFRFAYD--PKEAKNTALIFEKNVTGTSDIIFEN 318
T NADIRFNTKSESLLVK +YAGGARFRFAYD K + L FE+N+TGTSDIIFEN
Sbjct: 269 ATENADIRFNTKSESLLVKGYAGGARFRFAYDLQDKAPETPVLTFEQNITGTSDIIFEN 328

Query: 319 PIDDLKSLDGHQIIKVNGTADKHAFLRLSGKHQKGIYTL SLQRRPEGFLPKCKNAMIWRFM 378
PID+LKS L DGHQI+KVNGTADK+AFRLSGK++KGIYTL SLQ+RPEGF K +
Sbjct: 329 PIDNLKSLDGHQIVKVNGTADKNAFLRLSGKYRKGIYTL SLQRRPEGFFTKVQERDNIAIY 388

Query: 379 HNRLKPPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGKTAPVEGYRKGVQ 438
+ + ALRLN+KNSDIFDRTLPRKGLWLRVIDGHS+QWVQGKTAP+EGYRKGVQ
Sbjct: 389 AQQAQAANTLFALRLNDKNSDIFDRTLPRKGLWLRVIDGHSSQWVQGKTAPLEGYRKGVQ 448

Query: 439 LGGEVFTWQNESNQLSIGLMGGQAEQQRSTFHNPD TDNLTTGNVKGFGAGVYATWHQLQDK 498
LGGEVFTWQNESNQ S+GLMGGQAEQQRSTF NPDTDNLTTGN+KGFGAGVYATWHQLQDK
Sbjct: 449 LGGEVFTWQNESNQFSVGLMGGQAEQQRSTFRNPDTDNLTTGNMKGFGAGVYATWHQLQDK 508

Query: 499 QTGAYADSWMQYQRFRRHINTEDGTERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYL 558
QTGAY DSW+QYQRFRRH INTEDGTERFTSKGITASIEAGYNALLAEH T KG +R YL
Sbjct: 509 QTGAYIDSWVQYQRFRRHINTEDGTERFTSKGITASIEAGYNALLAEHVTGKGTQIRFYL 568

Query: 559 QPQAQLTYLGVNGKFSDSENHAVNLLGSRQLQTRVG VQAKAQFSLYKNIAIEPFPAAVNAL 618
QPQAQLTYLGVNGKFSDSEN+ VNLLG RQLQ+RVGVQAKAQF L KNI I+PFAAVN L
Sbjct: 569 QPQAQLTYLGVNGKFSDSENSQVNLLGPRQLQSRVG VQAKAQFLLNKNIVIQPFPAAVNTL 628

Query: 619 YHNKPFGEVMDGERRVINNKTAIESQLGVAVKIKSHL TLQATFNRQTGKHHQAKQGALNL 678
YH+KPFGEV+DGERRVINNKTAIESQ G+AVKIKSHL TLQATFNRQTGK H AKQGALNL
Sbjct: 629 YHSPKPFGEIDGERRVINNKTAIESQFGI AVKIKSHL TLQATFNRQTGKRHHAKQGALNL 688

Query: 679 QWTF 682
QWTF
Sbjct: 689 QWTF 692

tr Q93N94 Las (Fragment) [las] [Haemophilus influenzae biotype
aegyptius]

703
AA
align

Score = 663 bits (1710), Expect = 0.0

Identities = 337/480 (70%), Positives = 386/480 (80%), Gaps = 5/480 (1%)

```

Query: 204 IENYVSFAAHLYSGLDPFHYIEVTDNSHVIGQTISLDEFRLNSLWEPRWDSNVGKLKTT 263
      I+NY S      GL      + V + S V G T+ L + L++SLWEPR++S+V L+T
Sbjct: 228 IDNYRSKFPDETRGLT----LTVKNQSEVRGATLQLLKMVLQDSLWEPRFNSDVHHLETQ 283

Query: 264 NADIRFNTKSESLLVKEDYAG-GARFRFAYDPKEAKNTALIFEKNVTGTSDIIFENPIDD 322
      NA+IRFN+ + L V E+Y G G+RF ++PKEA L F+K+VTGTS+I+FE PIDD
Sbjct: 284 NANIRFNSTNTRLTVHENYQGDGSRFFIKFNPKEATQPVLTFDKDVGTGTSNIVFEKPIDD 343

Query: 323 LKSLDGHQIIKVNVTADKHAFLRLSGKHQKGIYTLSSLQQRPEGFLPKCKNAMIWRFMHNRL 382
      LKSLDGHQIIKVNVTADKHAFLRLSGKHQKGIYTLSSLQQRPEGF K + +
Sbjct: 344 LKSLDGHQIIKVNVTADKHAFLRLSGKHQKGIYTLSSLQQRPEGFFTKVQERDDIAIYAQQA 403

Query: 383 KPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGKTAPVEGYRKGVLGGGE 442
      + ALRLN+KNSDIFDRTLPRKGLWLR+I GH +Q VQGKTAPVEG RKG+QLGG+
Sbjct: 404 QAANTLFAFLRLNDKNSDIFDRTLPRKGLWLRLLISGHLSQDVQGKTAPVEGNRKGILGGD 463

Query: 443 VFTWQNESNQLSIGLMGGQAEQQRSTFHNPDNDNLTTGNVKGFGAGVYATWHQLQDKQTGA 502
      VF+ QN+ QLS GLMGGQAEQQRSTF N DTDNLTTG++KGFGAG+YATWHQLQDKQTGA
Sbjct: 464 VFSLQNQDYQLSFGLMGGQAEQQRSTFRNSDNDNLTTGSMKGFGAGIYATWHQLQDKQTGA 523

Query: 503 YADSWMQYQRFHRHINTEDGTERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQA 562
      Y DSW+YQYQRFHRHINTEDG ERFTSKGITASIEAGYNALLAEH T+KG +R YLQPQA
Sbjct: 524 YVDSWVQYQRFHRHINTEDGIERFTSKGITASIEAGYNALLAEHLTEKGTQIRFYLQPQA 583

Query: 563 QLTYLGVNGKFSSENHVNLLGSRQLQTRVGVAQAKAQFSLYKNIAIEPFAAVNALYHNK 622
      QLTYLGVNG +DS N+ VNLLGSRQLQ+RVG QAKAQF+ + +PF A+N++Y K
Sbjct: 584 QLTYLGVNGDLTDSGNSKVNLLGSRQLQSRVGAQAKAQFTFTNGVIFQPFVALNSIYQQK 643

Query: 623 PFGVEMDGERRVINNKTAIESQLGVAVKIKSHLTLOATFNROTGKHHQAKQGALNLQWTF 682
      PFGVE+D ERRVINNKTAIESQLGVA+KIKSHLTLOATFNROTGKHH AKQGALNLQWTF
Sbjct: 644 PFGVEIDNERRVINNKTAIESQLGVALKIKSHLTLOATFNROTGKHHHAKQGALNLQWTF 703

```

tr Q9JPL5 Autotransporter A [auta] [Neisseria meningitidis] 619 AA
align

Score = 389 bits (1000), Expect = e-107

Identities = 207/471 (43%), Positives = 302/471 (63%), Gaps = 24/471 (5%)

```

Query: 223 YIEVTDNSHVIGQTISLDEFRLNSLWEPRWDSNVGKLKTT-NADIRFNTKSESLLVKED 281
      ++++ + SHV GQ + L + L++SLWEPR S++ L+T+ NA IR NTK E L V +D
Sbjct: 162 HVKIENKSHVAGQVLELTKMTLKDSLWEPRRHSIDIHTLETSNARIRLNTKDEKLTVHKD 221

Query: 282 YAGGARFRFAYDPKEAKNTALIFEKNVTGTSDIIFENPIDDLKSLDGHQIIKVNVTADKH 341
      YAGGA F F YD +E+ AL FE V+G S ++ E ++LK+LDG ++I TAD
Sbjct: 222 YAGGADFLFGYDVRESDEPALTFEDKVSGQSGVVLERRPENLKTLDGRKLIAAK-TADSG 280

Query: 342 AFRLSGKHQKGIYTLSSLQQRPEGF-----LPKCKNAMIWR--FMHNRLKPPIPYCAL 391
      + +++G+Y L L+Q GF +P+ + + + + N L L
Sbjct: 281 SLAFKQNYRQGLYELLKQCEGGFCLGVQRLAIPAEAVLYAQQAYAANTL-----FGL 334

Query: 392 RLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGKTAPVEGYRKGVLGGGEVFTWQNESN 451
      R ++ D++ R+ LWLR I G S+Q ++G A +G+RKG VQ+GGEVF QNE +
Sbjct: 335 RAADRGGDVYAADPSRQKLWLRFIGGRSHQNIRGGAA-ADGWRKGVQIGGEVFVRQNEGS 393

```

Query: 452 QLSIGLMGGQAEQRSTFHNPDNDLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQ 511
+L+IG+MGG+A Q ++ + ++ G+G GVYA WHQL+DKQTGAY D W+QYQ
Sbjct: 394 RLAIGVMGGRAGQHASVNGKG--GAAGSDLYGYGGGVYAAWHQLRDKQTGAYLDGWLQYQ 451

Query: 512 RFRHRINTEDGTERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNG 571
RF+HRIN E+ ER+ +KG TAS+E GYNAL+AE KGN++R YLQPQAQ TYLGVNG
Sbjct: 452 RFKHRINDENRAERYKTKGWTASVEGGYNALVAEGIVGKGNVRFYLQPQAQFTYLGVNG 511

Query: 572 KFSDSENAHVNLLGSRQLQTRVGVQAKAQFSLYKNIAIEPFPAAVNALYHNKPFGEVMDGE 631
F+DSE V LLGS Q Q+R G++AK +F+L + ++PFAA N L+ +K FGVEVMDGE
Sbjct: 512 GFTDSEGTAVGLLGSQWQSRAGIRAKTRFALRNGVNLQPFPAFNVLRHSKSFGEVMDGE 571

Query: 632 RRVINNKTALIESQLGVAVKIKSHLTLQATFNROTGKHHQAKQALNLQWTF 682
++ + +TA+E + G+ K H++ + + ++T K+ AL+L+W F
Sbjct: 572 KQTLAGRTALEGRFGIEAGWKGHMSARIGYGKRTDGD---KEAALSLKWL 619

tr Q9ZF58 Virulence-associated protein VapA [vapA] [Neisseria
meningitidis]

619
AA
align

Score = 388 bits (997), Expect = e-106

Identities = 224/576 (38%), Positives = 338/576 (57%), Gaps = 46/576 (7%)

Query: 130 DTEEQIRKYFKECFNSNTK-----IRDYSTCQAEKFGSHPLIVK-SHIFSLGPKI---- 178
D ++KY + + N K ++D + E + + + ++I LGPK
Sbjct: 67 DNMPVVKKYITDTYGDNLKDAVKKQLQDLYKTRPEAWREENKKRTEEAYIEQLGPKFSILK 126

Query: 179 -KNSHINSEILSVGNYTEWANQVIHHIENYVSFAAHLYSGLDPFHYIEVTDNSHVIGQTI 237
KN + ++++ T +N + N + H +++ + SHV GQ +
Sbjct: 127 QKNPDLINKLVEDSVLTPHSNTSQTSLNNIFNKKLH-----VKIENKSHVAGQVL 176

Query: 238 SLDEFRLNSLWEPRWDSNVGKLT--NADIRFNTKSESLLVKEDYAGGARFRFAYDPKE 296
L + L++SLWEPR S++ L+T+ NA IR NTK E L V + Y GGA F F YD +E
Sbjct: 177 ELTKMTLKDSLWEPRRHSDIHTLETSDNARIRLNTKDEKLTVHKAYQGGADFLFGYDVRE 236

Query: 297 AKNTALIFEKNVTGTSDIIFENPIDDLKSLDGHQIIKVNGTADKHAFLSGKHQKGIYTL 356
+ AL FE+NV+G S+++ E ++LK+LDG ++I AD. ++F +++G+Y L
Sbjct: 237 SDEPALTFEQNVSGKSNVLERPENLKTLDGRKLIAAE-KADPNSFAFKQNYRQGLYEL 295

Query: 357 SLQQRPEGF-----LPKCKNAMIWR--FMHNRLKPPPIPYCALRLNNKNSDIFDRTL 406
L+Q GF +P+ + + + N L LR ++ D++
Sbjct: 296 LLKQCEGGFCLGVQRLAIPAEAVLYAQQAANTL-----FGLRAADRGGDVYAADPS 349

Query: 407 RKGLWLRVIDGHSNQWVQGTAPVEGYRKGVLGGEVFTWQNESNQLSIGLMGGQAEQRS 466
R+ LWLR I G S+Q ++G A +G RKGVL+GGEVF QNE ++L+IG+MGG+A Q +
Sbjct: 350 RQKLWLRFIGGRSHQNIIRGGAA-ADGRRKGVQIGGEVFVRQNEGSRLAIGVMGGRAGQHA 408

Query: 467 TFHNPDNDLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFHRINTEDGTERF 526
+ + + G+G GVYA WHQL+DKQTGAY D W+QYQRF+HRIN E+ ER+
Sbjct: 409 SVNGKG--GAAGSYLHGYGGGVYAAWHQLRDKQTGAYLDGWLQYQRFKHRINDENRVERY 466

Query: 527 TSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGKFSDSENAHVNLLGS 586
+KG TAS+E GYNAL+AE KGN++R YLQPQAQ TYLGVNG F+DSE V LLGS
Sbjct: 467 KTKGWTASVEGGYNALVAEGVVGKGNVRFYLQPQAQFTYLGVNGGFTDSEGTAVGLLGS 526

Query: 587 RQLQTRVGQAKAQFSLYKNIAIEPFAAVNALYHNKPFGEVMDGERRVINNKTAIESQLG 646
Q Q+R G++AK +F+L + ++PFAA N L+ +K FGVEMDGE++ + +TA+E + G
Sbjct: 527 GQWQSRAGIRAKTRFALRNGVNLQPFAAFNVLHRSKSFGEVMDGEKQTLAGRTALEGRFG 586

Query: 647 VAVKIKSHLTQLQATFNRTGKHHQAKQGALNLQWTF 682
+ K H+ + + ++T K+ AL+L+W F
Sbjct: 587 IEAGWKGHMFARIGYGKRTDGD---KEAALSLKWLF 619

tr Q9JSR2 Putative virulence associated protein [vapA] 619 AA
[Neisseria align
meningitidis (serogroup A)]

Score = 387 bits (993), Expect = e-106
Identities = 223/576 (38%), Positives = 337/576 (57%), Gaps = 46/576 (7%)

Query: 130 DTEEQIRKYFKECFNSNTK-----IRDYSTCQAEKFGSHPLIVK-SHIFSLGPKI---- 178
D ++KY + + N K ++D + E + + + ++I LGPK
Sbjct: 67 DNMPVVVKYITDITYGDNLDKDAVKKQLQDLYKTRPEAWENKKRTEEAYIEQLGPKFSILK 126

Query: 179 -KNSHINSEILSVGNYTEWANQVIHHIENYVSFAAHLYSGLDPFHYIEVTDNSHVIGQTI 237
KN + ++++ T +N + N + H +++ + SHV GQ +
Sbjct: 127 QKNPDLINKLVEDSVLTPHSNTSQTSLNNIFNKKLH-----VKIENKSHVAGQVL 176

Query: 238 SLDEFRLNSLWEPWRDSNVGKLT--NADIRFNTKSESLLVKEDYAGGARFRFAYDPKE 296
L + L++SLWEPR S++ L+T+ NA IR NTK E L V + Y GGA F F YD +E
Sbjct: 177 ELTKMTLKDSLWEPRRHSDIHMLETSNARIRLNTKDEKLTVHKAYQGGADFLFGYDVRE 236

Query: 297 AKNTALIFEKNVTGTSDIIFENPIDDLKSLDGHQIIKVNGTADKHAFLSGKHQKGIYTL 356
+ AL FE+ V+G S ++ E ++LK+LDG ++I AD ++F +++G+Y L
Sbjct: 237 SDKPALTFEEKVSGQSGVVLERRPENLKTLDGRKLIAAE-KADSNSFAFKQNYRQGLYEL 295

Query: 357 SLQQRPEGF-----LPKCKNAMIWR--FMHNRLKPPIPYCALRLNNKNSDIFDRTLP 406
L+Q GF +P+ + + + + N L LR ++ D++
Sbjct: 296 LLKQCEGGFCLGVQRLAIPAEAVLYAQYAAANTL-----FGLRAADRGDDVYAADPS 349

Query: 407 RKGLWLRVIDGHSNQWVQKGTAPVEGYRKGVLGGEVFTWQNESNQLSIGLMGGQAEQRS 466
R+ LWLR I G S+Q ++G A +G RKGVQ+GGEVF QNE ++L+IG+MGG+A Q +
Sbjct: 350 RQKLWLRFIGGRSHQNIRGGAA-ADGRRKG VQIGGEVFRVQNEGSRLAIGVMGGRAGQHA 408

Query: 467 TFHNPDTDNLTGTVGKFGAGVYATWHQLQDKQTGAYADSWMQYQFRHRINTEDGTERF 526
+ + + G+G GVYA WHQL+DKQTGAY D W+QYQRF+HRIN E+ ER+
Sbjct: 409 SVNGKG--GAAGSYLHGYGGGVYAAWHQLRDKQTGAYLDGWLQYQRFKHRINDENRAERY 466

Query: 527 TSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPPAQQLTYLGVNGKFSSENHVNLLGS 586
+KG TAS+E GYNAL+AE KGN++R YLQPPAQ TYLGVNG F+DSE V LLGS
Sbjct: 467 KTKGWTASVEGGYNALVAEGVVGKGNVRFYLQPPAQFTYLGVNGGFTDSEGTAVGLLGS 526

Query: 587 RQLQTRVGQAKAQFSLYKNIAIEPFAAVNALYHNKPFGEVMDGERRVINNKTAIESQLG 646
Q Q+R G++AK +F+L + ++PFAA N L+ +K FGVEMDGE++ + +TA+E + G
Sbjct: 527 GQWQSRAGIRAKTRFALRNGVNLQPFAAFNVLHRSKSFGEVMDGEKQTLAGRTALEGRFG 586

Query: 647 VAVKIKSHLTQLQATFNRTGKHHQAKQGALNLQWTF 682
+ K H++ + + ++T K+ AL+L+W F
Sbjct: 587 IEAGWKGHMSARIGYGKRTDGD---KEAALSLKWLF 619

tr Q6J5G0 Putative virulence-associated protein (Fragment) [Haemophilus 291 AA
influenzae]

align

Score = 219 bits (557), Expect = 2e-55

Identities = 136/291 (46%), Positives = 165/291 (55%), Gaps = 47/291 (16%)

Query: 235 QTISLDEFRLLENSLWEPRWDSNVGKLKTTNADIRFNTKSESLLVKE----- 280
Q I L E LENSLEPRWDS+V L N+ IRFNTK+ESL+V E

Sbjct: 1 QVIHLYELNLENSLWEPRWDS+VSYLSLYNSHIRFNTKNESESLVVGGENRIRPTPDNALETE 60

Query: 281 -----DYA-GGARFRFAYD--PKEAKNTALIFEKNVTGTS 312
DY G RFAYD +EA L + V G +

Sbjct: 61 KDFKSRFSNIGYHSSSKIISYNNRSDYVLGYPSIRFAYDLSEREADKPVLTLSKSKVRGKT 120

Query: 313 DIIFENP-IDDLKSLDGHQIIKVNGTADKHAFRLSGKHQKGIYTLSQLQRPEGFIPKCKN 371
I+FE +++LK+L Q+IK + AF L +++KG Y L LQQ P GF +

Sbjct: 121 AIVFEEKALNNLNLTQRQLIKTETDIEPDAFFLLEYYKKGRYRLFLQQCPNGFCIGVEK 180

Query: 372 AMIWRFM----HNRLKPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGKT 427
I + + ALRLN+KNSDIFDRTLPRKGLWLR+I GH +Q VQGKT

Sbjct: 181 LAIPTHLVASQAQAANTLFALRLNDKNSDIFDRTLPRKGLWLRISGHLSQDVQGKT 240

Query: 428 APVEGYRKGVLGGGEVFTWQNESQLSIGLMGGQAEQRSTFHNPDNDLTT 478

APVEG RKG+QLGG+VF+ QN+ Q S GLMGGQAEQRSTF NPDT N+TT

Sbjct: 241 APVEGNRKGVLGGDVFSLQNDYQFSFGLMGGQAEQRSTFRNPDTGNVTT 291

tr Q8X6C1 Putative beta-barrel outer membrane protein [z0402] 1349
[Escherichia coli
O157:H7] AA

align

Score = 142 bits (359), Expect = 2e-32

Identities = 100/327 (30%), Positives = 156/327 (47%), Gaps = 31/327 (9%)

Query: 371 NAMIWRFMHNRLKPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGK-TAP 429
N M +H RL N +D+ + +W+R GH N+W G

Sbjct: 1039 NTMFTTRLHERLG-----NTYYTDMVTGEQKQTTMWMRHEGGH-NKWRDGGGQLK 1087

Query: 430 VEGYRKGVLGGGEVFTW-QNESQLSIGLMGGQAEQRSTFHNPDNDLTTGNVKGFGAGV 488

+ R +QLGG+V W QN S++ +G+M G S + T +V G+ G+

Sbjct: 1088 TQSNRYVLQLGGDVAQWSQNGSDRWHVGMAGYGNDSKTISSRTGYRAKASVNGYSTGL 1147

Query: 489 YATWHQLQDKQTGAYADSWMQYQFRHRIINTED-GTERFTSKGITASIEAGYNALLAEHF 547

YATW+ + + GAY DSW QY F + + +D +E + SKG TAS+EAGY LAE

Sbjct: 1148 YATWYADDESNGAYLDSWAQYSWFDNTVKGDDLQSESYKSGFTASLEAGYKHKLAEFN 1207

Query: 548 TKKGNSLRVYLQPQAQLTYLGVNG-KFSDSENAHVNLGSRQLQTRVGVQA--KAQFSL- 603

+G Y+QPQAQ+T++GV K +S V+ G +QTR+GV+ K+ +

Sbjct: 1208 GSQGTRNEWYVQPQAQVTWMGVKADKHRESNGTLVHSNGDGNVQTRLGVKTWLKSHHKMD 1267

Query: 604 -YKNIAIEPFAAVNALYHNKPFVEM-----DGERRVINNKTAIESQLGVAVKIKSHL 655

K+ +PF VN L+++K F M DG R + KT +E QL + + ++

Sbjct: 1268 DGKSREFQPFVEVNWLHNSKDFSTSMGVSVTQDGARNIAEIKTGVEGQLNANLNVWGNV 1327

Query: 656 TLQATFNROTGKHHQAKQGALNLQWTF 682

+Q + + + ++W F

Sbjct: 1328 GVQV-----ADRGYNDSAMVGIKWQF 1349

tr Q7AH89 **AidA-I adhesin-like protein [ECs0362] [Escherichia coli** 1327
0157:H7] AA
align

Score = 142 bits (359), Expect = 2e-32

Identities = 100/327 (30%), Positives = 156/327 (47%), Gaps = 31/327 (9%)

Query: 371 NAMIWRFMHNRLKPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGK-TAP 429
N M +H RL N +D+ + +W+R GH N+W G

Sbjct: 1017 NTMFTTRLHERLG-----NTYYTDMVTGEQKQTTMWMRHEGGH-NKWRDGSQGLK 1065

Query: 430 VEGYRKGVQLGGEVFTW-QNESNQLSIGLMGGQAEQRSTFHNPDNDLTTGNVKGFGAGV 488

+ R +QLGG+V W QN S++ +G+M G S + T +V G+ G+

Sbjct: 1066 TQSNRYVLQLGGDVAQWSQNGSDRWHVGMAGYGNSSDKTSSRTGYRAKASVNGYSTGL 1125

Query: 489 YATWHQLQDKQTGAYADSWMQYQFRHRINTED-GTERFTSKGITASIEAGYNALLAEHF 547

YATW+ + + GAY DSW QY F + + +D +E + SKG TAS+EAGY LAE

Sbjct: 1126 YATWYADDESNGAYLDSWAQYSWFDNTVKGDDLQSESYKSKGFTASLEAGYKHKLAEFN 1185

Query: 548 TKKGNSLRVYLQPQAQLTYLGVNG-KFSDSENAHVNLLGSRQLQTRVGVQA--KAQFSL- 603

+G Y+QPQAQ+T++GV K +S V+ G +QTR+GV+ K+ +

Sbjct: 1186 GSQGTRNEWYVQPQAQVTWGMVKADKHRESNGTLVHSNGDGNVQTRLGVKTLWLSHHKMD 1245

Query: 604 -YKNIAIEPFAAVNALYHNKPFGEVEM-----DGERRVINNKTAIESQLGVAVKIKSHL 655

K+ +PF VN L+++K F M DG R + KT +E QL + + ++

Sbjct: 1246 DGKSREFQPFVEVNWLHNSKDFSTSMGVSVTQDGARNIAEIKTGVEGQLNANLNVWGNV 1305

Query: 656 TLQATFNROTGKHHQAKQGALNLQWTF 682

+Q + + + ++W F

Sbjct: 1306 GVQV-----ADRGYNDSAMVGIKWQF 1327

tr Q6KCW6 **YapH homolog [yapH] [Escherichia coli]** 2001 AA
align

Score = 139 bits (351), Expect = 1e-31

Identities = 87/294 (29%), Positives = 145/294 (48%), Gaps = 10/294 (3%)

Query: 398 SDIFDRTLPRKGLWLRVIDGHSNQWVQGKTA-PVEGYRKGVQLGGEVFTWQNESNQLSIG 456

+D+F +W+R I GH N+W + + R VQLGG + W + ++L G

Sbjct: 1709 TDVFTGEKKATSMWMRHIGGH-NRWKDSSSQLNTQSNRYVVQLGGSIAQWTDGQDRLQQG 1767

Query: 457 LMGGQAEQRSTFHNPDNDLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQFRHR 516

+M G ++S+ + + + G + G+ G+Y TW Q GAY D+W+QY F +

Sbjct: 1768 IMAGYGNEKSSTTSSLSGYKSKGAINGYSTGLYGTWQQNDGNDGAYVDTWIQYGFNFNT 1827

Query: 517 INTED-GTERFTSKGITASIEAGYNALLAEHF+TKKGNSLRVYLQPQAQLTYLGVN-GKFS 574

+N E E + S+G T S+EAGY E +G+ Y+QPQ+Q+T++ V + +

Sbjct: 1828 VNGEKLAAESWKS RGTGSVEAGYTFKAGEFTGSQGS HYDWYIQPQSQITWMNV RASEHT 1887

Query: 575 DSENAHV NLLGSRQLQTRVGV---QAKAQFSLYKNIAIEPFAAVNALYHNKPF GVEMDG 630
+ V L G +Q+R+GV + K+ K EPF VN +++ + +GV+MD

Sbjct: 1888 EKNGTKVQLSGDGNIQSRLGVRTYLKGKSASDDNKAHQFEPFVEVNW IHNTRSWGVKMDN 1947

Query: 631 ERRVINNKTAI-ESQLGVA VKIKSHLT LQATFNRTG-KHHQAKQGALNLQWTF 682
+ T I E + GV K+ +L + Q G K + Q L +++ F

Sbjct: 1948 TALSQDGATNIAEVKTGVQGLSDNLNVWGNVGVQAGDKGYSDAQAMLGIKYIF 2001

tr Q8FFF9 **Hypothetical protein ydeU [ydeU] [Escherichia coli** 291 AA
06] align

Score = 137 bits (344), Expect = 9e-31

Identities = 85/282 (30%), Positives = 141/282 (49%), Gaps = 10/282 (3%)

Query: 410 LWLRVIDGHSNQWVQGKTA-PVEGYRKG VQLGGEVFTWQNESNQLSIGLMGGQAEQRSTF 468
+W+R I GH N+W + + R VQLGG + W + ++L G+M G ++S+

Sbjct: 11 MWMRHIGGH-NRWKDSSSQLNTQSNRYVVQLGGSIAQWTDGQDRLQQGIMAGYGNEKSST 69

Query: 469 HNPDTDNLT TGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFRRHINTED-GTERFT 527
+ + + G + G+ G+Y TW Q GAY D+W+QY F + +N E E +

Sbjct: 70 TSSLSGYKSKGAINGYSTGLYGTWQQNDGNDNGAYVDTWIQYGFNNTVNGEKLAAESWK 129

Query: 528 SKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVN-GKFSDSENAHV NLLGS 586
S+G T S+EAGY E +G+ Y+QPQ+Q+T++ V + ++ V L G

Sbjct: 130 SRGTGSVEAGYTFKAGEFTGSQGS HYDWYIQPQSQITWMNV RASEHTEKNGTKVQLSGD 189

Query: 587 RQLQTRVGV----QAKAQFSLYKNIAIEPFAAVNALYHNKPF GVEMDGERRVINNKTAI- 641
+Q+R+GV + K+ K EPF VN +++ + +GV+MD + T I

Sbjct: 190 GNIQSRLGVRTYLKGKSASDDNKAHQFEPFVEVNW IHNTRSWGVKMDNTALSQDGATNIA 249

Query: 642 ESQLGVA VKIKSHLT LQATFNRTG-KHHQAKQGALNLQWTF 682
E + GV K+ +L + Q G K + Q L +++ F

Sbjct: 250 EVKTGVQGLSDNLNVWGNVGVQAGDKGYSDAQAMLGIKYIF 291

tr Q7BCK4 **VirG [virG] [Shigella flexneri]** 1102 AA
align

Score = 134 bits (338), Expect = 4e-30

Identities = 88/297 (29%), Positives = 149/297 (49%), Gaps = 12/297 (4%)

Query: 398 SDIFDRT---LPRKGLWLRVIDGHSNQWVQGKTAPVEGYRKG VQLGGEVF-TWQNESNQL 453
S I D T L +W+R + GH+ + + R Q+GG++ T + + L

Sbjct: 806 STIVDPTTGQLSETTMWIRTVG GHNEHNADRQLKTTANRMVYQIGDILKTNFTDHDGL 865

Query: 454 SIGLMGGQAEQRSTFHNPD TGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRF 513
+G+MG Q S HN T + G V G+ AG+Y++W Q + ++TG Y D+W+QY F

Sbjct: 866 HVGIMGAYGYQDSKTHNKYTSYSSRGTVSGYTAGLYSSWFQDEKERTGLYMDAWLQYSWF 925

Query: 514 RHRINTEDGT-ERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLG VNGK 572
+ + + T E+++SKGIT ++EAGY +YL PQ Q+T GV

Sbjct: 926 NNTVKGDGLTGEKYSSKGITGALEAGYIYPTIRWTAHNNIDNALYLN PQVQITRHGVKAN 985

Query: 573 -FSDSENAHVNLLGSRQLQTRVGVOAK--AQFSLYKNI--AIEPFAAVNALYHNKPFGVE 627
 + + V G +Q ++G++ +Q + K EPF VN + +K +GV
 Sbjct: 986 DYIEHNGTMVTSSGGNNIQA KLGLRTSLISQSCIDKETLRKFEPFLEVNWKSSKQYGV I 1045

Query: 628 MDG-ERRVINNKTAIESQLGVAVKIKSHLT LQATFNRQTGKH-HQAKQGALNLQWTF 682
 M+G I N+ IE + GV ++ +L++ ++Q G + ++ QG L +++TF
 Sbjct: 1046 MNGMSNHQIGNRNVIELKTGVGGRLADNLSIWGNVSQQLGNNSYRDTQGILGVKYTF 1102

tr Q99Q93 IcsA (VirG), outermembrane protein exposed to the bacterial surface 1102
 by a C-terminal autotransporter domain and involved in the movement of intracellular bacteria by binding to N-WASP (Invasion protein) [icsA (virG)] [Shigella flexneri] AA
align

Score = 134 bits (338), Expect = 4e-30
 Identities = 88/297 (29%), Positives = 149/297 (49%), Gaps = 12/297 (4%)

Query: 398 SDIFDRT--LPRKGLWLRVIDGHSNQWVQGKTAPVEGYRKG VQLGGEVF-TWQNESNQL 453
 S I D T L +W+R + GH+ + + R Q+GG++ T + + L
 Sbjct: 806 STIVDPTTGQLSETTMWIRTVG GHNEHN LADRQLKTTANRMVYQIGGDILKTNFTDHDGL 865

Query: 454 SIGLMGGQAEQRSTFHNPDTDNLT TGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRF 513
 +G+MG Q S HN T + G V G+ AG+Y++W Q + ++TG Y D+W+QY F
 Sbjct: 866 HVGIMGAYGYQDSKTHNKYTSYSSRGTVSGYTAGLYSSWFQDEKERTGLYMDAWLQYSWF 925

Query: 514 RHRINTEDGT-ERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGK 572
 + + + T E+++SKGIT ++EAGY +YL PQ Q+T GV
 Sbjct: 926 NNTVKGDGLTGEKYSSKGITGALEAGYIYPTIRWTAHNNIDNALYLN PQVQITRHGVKAN 985

Query: 573 -FSDSENAHVNLLGSRQLQTRVGVOAK--AQFSLYKNI--AIEPFAAVNALYHNKPFGVE 627
 + + V G +Q ++G++ +Q + K EPF VN + +K +GV
 Sbjct: 986 DYIEHNGTMVTSSGGNNIQA KLGLRTSLISQSCIDKETLRKFEPFLEVNWKSSKQYGV I 1045

Query: 628 MDG-ERRVINNKTAIESQLGVAVKIKSHLT LQATFNRQTGKH-HQAKQGALNLQWTF 682
 M+G I N+ IE + GV ++ +L++ ++Q G + ++ QG L +++TF
 Sbjct: 1046 MNGMSNHQIGNRNVIELKTGVGGRLADNLSIWGNVSQQLGNNSYRDTQGILGVKYTF 1102

tr Q52298 Virulence-associated VirG [Plasmid pMYSH6000] 1102 AA
align

Score = 134 bits (338), Expect = 4e-30
 Identities = 88/297 (29%), Positives = 149/297 (49%), Gaps = 12/297 (4%)

Query: 398 SDIFDRT--LPRKGLWLRVIDGHSNQWVQGKTAPVEGYRKG VQLGGEVF-TWQNESNQL 453
 S I D T L +W+R + GH+ + + R Q+GG++ T + + L
 Sbjct: 806 STIVDPTTGQLSETTMWIRTVG GHNEHN LADRQLKTTANRMVYQIGGDILKTNFTDHDGL 865

Query: 454 SIGLMGGQAEQRSTFHNPDTDNLT TGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRF 513
 +G+MG Q S HN T + G V G+ AG+Y++W Q + ++TG Y D+W+QY F
 Sbjct: 866 HVGIMGAYGYQDSKTHNKYTSYSSRGTVSGYTAGLYSSWFQDEKERTGLYMDAWLQYSWF 925

Query: 514 RHRINTEDGT-ERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGK 572
 + + + T E+++SKGIT ++EAGY +YL PQ Q+T GV
 Sbjct: 926 NNTVKGDGLTGEKYSSKGITGALEAGYIYPTIRWTAHNNIDNALYLNPPVQITRHGVKAN 985

Query: 573 -FSDSENAHVNLGSRQLQTRVGQAK--AQFSLYKNI--AIEPFAAVNALYHNKPFGVE 627
 + + V G +Q ++G++ +Q + K EPF VN + +K +GV
 Sbjct: 986 DYIEHNGTMVTSSGGNNIQAQKLGLRTSLISQSCIDKETLRKFEPFLEVNWKWSSQYGVVI 1045

Query: 628 MDG-ERRVINNKTAIESQLGVAVKIKSHLTLQATFNRTGKH-HQAKQGALNLQWTF 682
 M+G I N+ IE + GV ++ +L++ ++Q G + ++ QG L +++TF
 Sbjct: 1046 MNGMSNHQIGNRVIELKTGVGGRLADNLSIWGNVSQQLGNNSYRDTQGILGVKYTF 1102

tr Q9JMS3 YchA protein [ychA] [Escherichia coli] 1371 AA
align

Score = 134 bits (337), Expect = 6e-30

Identities = 122/461 (26%), Positives = 193/461 (41%), Gaps = 46/461 (9%)

Query: 266 DIRFNTKSESLLVKEDYAGG---ARFRFAYDPKEAKNTALIFEKNVTGTSDIIFENPIDD 322
 D+R L + +Y G D + L+ + + +G + + N
 Sbjct: 913 DLRSAAPGNILTIGGNYTGNNGTLLINTVLDDSSSATDKLVIKGDASGKTRVAVTNVGGS 972

Query: 323 -LKSLDGHQIIKVNGTADKHAFRLSGKHQKGIYTLSSLQQRPEG-----FLPKCKNAMIWR 376
 +L+ ++I V+G A F +G+ G Y +L + P +L KN R
 Sbjct: 973 GANTLNSIEVIHVDGNAANAEFIQAGRIAAGAYDYLGRGPGSNYGNWYLSSSKNTPEPR 1032

Query: 377 -----FMHNRLKPPPIPYCALRLNNKNSDIFDRTLPRKG-----LWL 412
 N L+P + N+ R R G +W+
 Sbjct: 1033 PDPEPTPEGHDNNLRPEASSYTANIAAANTMFVTRLHERLGQTQYVDAITGEPKATSMWM 1092

Query: 413 RVIDGHSNQWVQ GK-TAPVEGYRKGVQLGGEVFTWQ-NESNQLSIGLMGGQAEQRSTFHN 470
 R GH N+W G + R +QLGG++ W +N+ +G+M G S+
 Sbjct: 1093 RHEGGH-NRWRDGSQGLKTQSNRYVIQLGGDIAQWDWGGTNRWHLGVMAGYGNNHSSTGA 1151

Query: 471 PDTDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQFRHRINTEDGT--ERFTS 528
 T + G+V G+ G+YATW+ + GAY D+W QY F + + DG E + S
 Sbjct: 1152 VRTGYHSGKSVNGYSTGLYATWYADDETHNGAYLDTWAQYGWFDNHVK-GDGLPGESWKS 1210

Query: 529 KGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNG-KFSDSENAHVNLGSR 587
 KG+TAS+E GY + E + GN Y+QPQAQL ++GV + +S + G
 Sbjct: 1211 KGLTASLETGYAWKIGEFSSNYGNLNEWYVQPQAQLVWMGVKADELYESNGTLIESTGDG 1270

Query: 588 QLQTRVGVA---KAQFSLYKNIAIEPFAAVNALYHNKPFGVEMDGERRVINNKTAI-E 642
 + TR+GV+ + K+ PF VN L++ + FGV M+GE + I E
 Sbjct: 1271 NVHTRLGVKTIKRLNKMDDGKSREFSPFVEVNLHNTRDFGVRMNGEPVYQDGTNRNIGE 1330

Query: 643 SQLGVAVKIKSHLTLQATFNRTG-KHHQAKQGALNLQWTF 682
 + GV +I HL L Q G K + L +++TF
 Sbjct: 1331 VKTGVEGQINPHLNLWGNVRVQVGDKGYNDTSAMLGVKYTF 1371

tr Q6WE27 IcsA [icsA] [Shigella flexneri] 1102 AA
align

9/14/04

Sbjct: 3624 KVS---GPAFKPFVEANWIHNTKDFGTTLDGVTVKQAGTANIAELKLGVDGQVNSQLNLW 3680

Query: 659 ATFNRQTG-KHHQAKQGALNLQWTF 682

+Q G K + L +++ F

Sbjct: 3681 GNIGQQVGNKGYSETSVVLGVKYNF 3705

tr Q8CZU2 Putative autotransporter adhesin [yapH] [Yersinia pestis] 3710 AA
align

Score = 130 bits (326), Expect = 1e-28

Identities = 106/385 (27%), Positives = 170/385 (43%), Gaps = 36/385 (9%)

Query: 324 KSLDGHQIIKVN GTADKHAFRLSGKHQKGIYTL SLQQRPEGFLPKCKNAMIWRFMHNRLK 383

++ +G +II V G + F L+G+ G Y L Q P N + +R

Sbjct: 3336 RTFEGIKIIDVGGDSAGQ-FTLN GRAVGGAYEYFLYQGGAS-TPDDGNWYLRTEADDRRP 3393

Query: 384 PPIPYCALRLNNKN-----SDIFDRTLPRKGLWLRVIDGHSNQWVQGKTA 428

P Y A N +D+F LWLR H+

Sbjct: 3394 EPASYTANLAAANNMFV TSLADRMGETLYTDVFTGEQKTTSLWLRNEGSHNRSRDDSGEL 3453

Query: 429 PVEGYRKGVQLGGEVFTWQNESNQL-SIGLMGGQAEQRSTFHNPD TDNLTTGNVKGFGAG 487

+ R +QLGG+V W + L +G+M G A S+ +TG+V G+ G

Sbjct: 3454 KTQDNRYVMQLGGDVAQWSRNAQDLWRVGVMMAGYANSSSSTVAQVAGYRSTGSVDGYSVG 3513

Query: 488 VYATWHQLQDKQTGAYADSWMQYQRFRRHINTED-GTERFTSKGITASIEAGYNALLAEH 546

+Y +W TGAY DSW+QY F +R++ +D TE++ SKG TAS+E GY + E

Sbjct: 3514 IYGSWLADNADDTGAYVDSWVQYSWFDNRVSGQDLATEKYDSKGFTASVEGGYAFKVGES 3573

Query: 547 FTKKGNLSLRVYLQPPAQLTYLGVNG-KFSDSENAHVNLLGSRQLQTRVGVQA-----KA 599

+ ++QP+AQ+ ++GV +++ ++ G+ +QTR+G +A KA

Sbjct: 3574 VNQS-----YFIQPKAQVVMGVKADDHTETNGTVISGDGNGNIQTRLGAKAFINPSDKA 3628

Query: 600 QFSLYKNIAIEPFAAVNALYHNKPFVEMDGER-RVINNKTAIESQLGVAVKIKSHLT LQ 658

+ S A +PF N +++ K FG +DG + E +LGV ++ S L L

Sbjct: 3629 KVS---GPAFKPFVEANWIHNTKDFGTTLDGVTVKQAGTANIAELKLGVDGQVNSQLNLW 3685

Query: 659 ATFNRQTG-KHHQAKQGALNLQWTF 682

+Q G K + L +++ F

Sbjct: 3686 GNIGQQVGNKGYSETSVVLGVKYNF 3710

tr Q9F289 YapD protein [yapD] [Yersinia pestis] 1457 AA
align

Score = 130 bits (326), Expect = 1e-28

Identities = 106/385 (27%), Positives = 170/385 (43%), Gaps = 36/385 (9%)

Query: 324 KSLDGHQIIKVN GTADKHAFRLSGKHQKGIYTL SLQQRPEGFLPKCKNAMIWRFMHNRLK 383

++ +G +II V G + F L+G+ G Y L Q P N + +R

Sbjct: 1083 RTFEGIKIIDVGGDSAGQ-FTLN GRAVGGAYEYFLYQGGAS-TPDDGNWYLRTEADDRRP 1140

Query: 384 PPIPYCALRLNNKN-----SDIFDRTLPRKGLWLRVIDGHSNQWVQGKTA 428

P Y A N +D+F LWLR H+

Sbjct: 1141 EPASYTANLAAANNMFV TSLADRMGETLYTDVFTGEQKTTSLWLRNEGSHNRSRDDSGEL 1200

Query: 429 PVEGYRKGVQLGGEVFTWQNESNQL-SIGLMGGQAEQRSTFHNPDNDNLTGNNVKGFGAG 487
+ R +QLGG+V W + L +G+M G A S+ +TG+V G+ G
Sbjct: 1201 KTQDNRYVMQLGGDVAQWSRNAQDLWRVGVVMAGYANSSSSTVAQVAGYRSTGSVDGYSVG 1260

Query: 488 VYATWHQLQDKQTGAYADSWMQYQFRHRINTED-GTERFTSKGITASIEAGYNALLAEH 546
+Y +W TGAY DSW+QY F +R++ +D TE++ SKG TAS+E GY + E
Sbjct: 1261 IYGSWLADNADDTGAYVDSWVQYSWFDNRVSGQDLATEKYDSKGFTASVEGGYAFKVGES 1320

Query: 547 FTKKGNLSRVYLQPQAQLTYLGVNG-KFSDSENAHVNLLGSRQLQTRVGVQA-----KA 599
+ ++QP+AQ+ ++GV +++ ++ G+ +QTR+G +A KA
Sbjct: 1321 VNQS-----YFIQPKAQVVMGVKADDHTETNGTVISGDGNGNIQTRLGAKAFINPSDKA 1375

Query: 600 QFSLYKNIAIEPFPAVNALYHNKPFVEMDGER-RVINNKTAIESQLGVAVKIKSHLTLQ 658
+ S A +PF N +++ K FG +DG + E +LGV ++ S L L
Sbjct: 1376 KVS---GPAFKPFVEANWIHNTKDFGTTLDGVTVKQAGTANIAELKLGVDGQVNSQLNLW 1432

Query: 659 ATFNRQTG-KHHQAKQGALNLQWTF 682
+Q G K + L +++ F
Sbjct: 1433 GNIGQQVGNKGYSETSVVLGVKYNF 1457

tr Q9F285 YapH protein [yapH] [Yersinia pestis] 3705 AA
align

Score = 130 bits (326), Expect = 1e-28
Identities = 106/385 (27%), Positives = 170/385 (43%), Gaps = 36/385 (9%)

Query: 324 KSLDGHQIIKVGNTADKHAFLSGKHQKGIYTLSQLQRPEGFLPKCKNAMIWRFMHNRLK 383
++ +G +II V G + F L+G+ G Y L Q P N + +R
Sbjct: 3331 RTFEGIKIIDVGDSAGQ-FTLNGRAVGGAYEYFLYQGGAS-TPDDGNWYLRTTEADRRP 3388

Query: 384 PPIPYCALRLNNKN-----SDIFDRTLPRKGLWLRVIDGHSNQWVQGKTA 428
P Y A N +D+F LWLR H+
Sbjct: 3389 EPASYTANLAAANNMFVTSLADRMGETLYTDVFTGEQKTTSLWLRNEGSHNRSRDDSGEL 3448

Query: 429 PVEGYRKGVQLGGEVFTWQNESNQL-SIGLMGGQAEQRSTFHNPDNDNLTGNNVKGFGAG 487
+ R +QLGG+V W + L +G+M G A S+ +TG+V G+ G
Sbjct: 3449 KTQDNRYVMQLGGDVAQWSRNAQDLWRVGVVMAGYANSSSSTVAQVAGYRSTGSVDGYSVG 3508

Query: 488 VYATWHQLQDKQTGAYADSWMQYQFRHRINTED-GTERFTSKGITASIEAGYNALLAEH 546
+Y +W TGAY DSW+QY F +R++ +D TE++ SKG TAS+E GY + E
Sbjct: 3509 IYGSWLADNADDTGAYVDSWVQYSWFDNRVSGQDLATEKYDSKGFTASVEGGYAFKVGES 3568

Query: 547 FTKKGNLSRVYLQPQAQLTYLGVNG-KFSDSENAHVNLLGSRQLQTRVGVQA-----KA 599
+ ++QP+AQ+ ++GV +++ ++ G+ +QTR+G +A KA
Sbjct: 3569 VNQS-----YFIQPKAQVVMGVKADDHTETNGTVISGDGNGNIQTRLGAKAFINPSDKA 3623

Query: 600 QFSLYKNIAIEPFPAVNALYHNKPFVEMDGER-RVINNKTAIESQLGVAVKIKSHLTLQ 658
+ S A +PF N +++ K FG +DG + E +LGV ++ S L L
Sbjct: 3624 KVS---GPAFKPFVEANWIHNTKDFGTTLDGVTVKQAGTANIAELKLGVDGQVNSQLNLW 3680

Query: 659 ATFNRQTG-KHHQAKQGALNLQWTF 682
+Q G K + L +++ F
Sbjct: 3681 GNIGQQVGNKGYSETSVVLGVKYNF 3705

tr Q74QP7 Putative autotransporter protein [yapH] [Yersinia pestis] 3710 AA
align

Score = 130 bits (326), Expect = 1e-28

Identities = 106/385 (27%), Positives = 170/385 (43%), Gaps = 36/385 (9%)

Query: 324 KSLDGHQIIKVNGTADKHAFRLSGKHQKGIYTLSQLQRPEGLPKCKNAMIWRFMHNRLK 383
++ +G +II V G + F L+G+ G Y L Q P N + +R
Sbjct: 3336 RTFEGIKIIDVGGDSAGQ-FTLNGRAVGGAYEYFLYQGGAS-TPDDGNWYLRTTEADRRP 3393

Query: 384 PPIPYCALRLNKN-----SDIFDRTLPRKGLWLRVIDGHSNQWVQGKTA 428
P Y A N +D+F LWLR H+
Sbjct: 3394 EPASYTANLAAANNMFVTSLADRMGETLYTDVFTGEQKTTSLWLRNEGSHNRSRDDSGEL 3453

Query: 429 PVEGYRKGVQLGGEVFTWQNESNQL-SIGLMGGQAEQRSTFHNPDNDLTGNNKGVGAG 487
+ R +QLGG+V W + L +G+M G A S+ +TG+V G+ G
Sbjct: 3454 KTQDNRYVMQLGGDVAQWSRNAQDLWRVGMAGYANSSSSTVAQVAGYRSTGSDVDGYSVG 3513

Query: 488 VYATWHQLQDKQGTGAYADSWMQYQFRHRINTED-GTERFTSKGITASIEAGYNALLAEH 546
+Y +W TGAY DSW+QY F +R++ +D TE++ SKG TAS+E GY + E
Sbjct: 3514 IYGSWLADNADDTGAYVDSWVQYSWFDNRVSGQDLATEKYDSKGFTASVEGGYAFKVGES 3573

Query: 547 FTKKGNSLRVYLQPPAQLTYLGVNG-KFSDSENAHVNLLGSRQLQTRVGVQA-----KA 599
+ ++QP+AQ+ ++GV +++ ++ G+ +QTR+G +A KA
Sbjct: 3574 VNQS-----YFIQPKAQVVMGMVKADDHTETNGTVISGDGNGNIQTRLGAKAFINPSDKA 3628

Query: 600 QFSLYKNIAIEPFAAVNALYHNKPFGVEMDGER-RVINNKTAIESQLGVAVKIKSHLTLQ 658
+ S A +PF N +++ K FG +DG + E +LGV ++ S L L
Sbjct: 3629 KVS---GPAFKPFVEANWIHNTKDFGTTLDGVTVKQAGTANIAELKLGVDGQVNSQLNLW 3685

Query: 659 ATFNRTG-KHHQAKQGALNLQWTF 682
+Q G K + L +++ F
Sbjct: 3686 GNIGQQVGNKGYSETSVVLGVKYNF 3710

tr Q9F286 YapG protein (Putative ATP-binding transport component) 994
(Putative AA
autotransporter protein) [yapG] [Yersinia pestis] align

Score = 129 bits (324), Expect = 2e-28

Identities = 84/272 (30%), Positives = 132/272 (47%), Gaps = 12/272 (4%)

Query: 410 LWLRVIDGHSNQWVQGKTAPVEGYRKGVQLGGEVFTWQ-NESNQLSIGLMGGQAEQRSTF 468
+W+ G S + ++G VQLGG++ W NE IG++ G + +
Sbjct: 723 MWMHTQGGRSQFGHTVEQLNIKGNYYSVQLGGDIAQWATNEQSGRIGMLAGLGKATNHS 782

Query: 469 HNPDTNDLTGNNKGVGAGVYATWHQLQDKQGTGAYADSWMQYQFRHRINTED-GTERFT 527
H+ T + G V G+ G+YATW Q TG Y D+ QY F + +N +D E++
Sbjct: 783 HSKVTSYHSRGAVDGYNLGIYATWFADQQHNTGVYIDTLAQYSWFNNVAVNGQDKAEKYYK 842

Query: 528 SKGITASIEAGYNALLAEHFTTKKGNSLRVYLQPPAQLTYLGVNGK-FSDSENAHVNLLGS 586
S G T SIE+GY LA + L ++QP AQ+T+ G+N + ++ A V+ +
Sbjct: 843 SSGFTTSIESGYTFNLA-----NSDQLSYFIQPNQITWAGINAQTHKTADGAVVSYRNN 897

Query: 587 RQLQTRVGVQAKAQFSLYKNIAIEPFAAVNALYHNKPFGVEMDGERRVINNK--TAIESQ 644

TR+G +A Q N PF AVN ++ N+ G + G+ I+NK + E
 Sbjct: 898 GHFITRIGAKAYLQTHDTLNTKFTPFVAVNWIHQNTGTTISGQG--IDNKIQNSTEFN 955
 Query: 645 LGVAVKIKSHLTLQATFNRQTGKHHQAKQGAL 676
 +GV +I L + A N Q G+++ AL
 Sbjct: 956 VGVESQIDQQLHIWANINHQIGRYNYTDTNAL 987

tr Q74RV9 Putative autotransporter protein [YP2907] [Yersinia pestis] 994 AA

align

Score = 129 bits (324), Expect = 2e-28

Identities = 84/272 (30%), Positives = 132/272 (47%), Gaps = 12/272 (4%)

Query: 410 LWLRVIDGHSNQWVQGKTAPVEGYRKGVLGGEVFTWQ-NESNQLSIGLMGGQAEQRSTF 468
 +W+ G S + ++G VQLGG++ W NE IG++ G + +
 Sbjct: 723 MWMHTQGGRSQFGHTVEQLNIKNYYSVQLGGDIAQWATNEQSGRIGMLAGLGKATNHS 782
 Query: 469 HNPDTDNLTGTVKGFAGVYATWHQLQDKQTGAYADSWMQYQFRHRINTED-GTERFT 527
 H+ T + G V G+ G+YATW Q TG Y D+ QY F + +N +D E++
 Sbjct: 783 HSKVTSYHSRGAVDGYNLGIYATWFADQQHNTGVYIDTLAQYSWFNNAVNGQDKAEKEYK 842
 Query: 528 SKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGK-FSDSENAHVNLLGS 586
 S G T SIE+GY LA + L ++QP AQ+T+ G+N + ++ A V+ +
 Sbjct: 843 SSGFTTSIESGYTFNLA-----NSDQLSYFIQPNQAQITWAGINAQTHKTADGAVVSYRNN 897
 Query: 587 RQLQTRVGQAKAQFSLYKNIAIEPFAAVNALYHNKPFGEVMDGERRVINNK--TAIESQ 644
 TR+G +A Q N PF AVN ++ N+ G + G+ I+NK + E
 Sbjct: 898 GHFITRIGAKAYLQTHDTLNTKFTPFVAVNWIHQNTGTTISGQG--IDNKIQNSTEFN 955
 Query: 645 LGVAVKIKSHLTLQATFNRQTGKHHQAKQGAL 676
 +GV +I L + A N Q G+++ AL
 Sbjct: 956 VGVESQIDQQLHIWANINHQIGRYNYTDTNAL 987

sp P77286 Hypothetical protein ydeU [ydeU] [Escherichia 466 AA
 YDEU_ECOLI coli] align

Score = 129 bits (323), Expect = 2e-28

Identities = 87/323 (26%), Positives = 155/323 (47%), Gaps = 23/323 (7%)

Query: 371 NAMIWRFMHNRLKPPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGKTAPV 430
 N++ +H+RL P + +D +W+R + GH
 Sbjct: 156 NSLFSHRLHDLRGEP-----QYTDLSLHSGSASSMWMRHVGGHERSRAGDQQLNT 205
 Query: 431 EGYRKGVLGGEVFTWQ-NESNQLSIGLMGGQAEQRSTFHNPDNDLTGTVKGFAGVY 489
 + R +QLGG++ W N ++ +G+M G A Q S + + G + G+ AG+Y
 Sbjct: 206 QANRYVLQLGGDLAQWSSNAQDRWHLGVMAGYANQHSNTQSNRVGYKSDGRISGYSAGLY 265
 Query: 490 ATWHQLQDKQTGAYADSWMQYQFRHRINTED-GTERFTSKGITASIEAGYNALLAEHFT 548
 ATW+Q +TGAY DSW Y F + +++++ + + S+G+TAS+E GY
 Sbjct: 266 ATWYQNDANKTGAYVDSWALYNWFDNSVSSDNRSADDYDSRGVTASVEGGYTFEAGTFSG 325
 Query: 549 KKGNSLRVYLQPQAQLTYLGV-NGKFSDSENAHVNLLGSRQLQTRVGQ----AKAQFSL 603

```

      +G      Y+QPQAQ+T++GV +   +   +   G   +QTR+GV+   +   Q
Sbjct: 326 SEGTLNTWYVQPQAQITWMGVKDSHDTRKDGTRIETEGDGNVQTRLGVKTYLNSHHQRDD 385

Query: 604 YKNIAIEPFAAVNALYHNKPFGEVMDGE---RRVINNKTAIESQLGVAVKIKSHLTLQAT 660
      K   +P+   N + ++K + V+M+G+   R   N   E + GV K+ ++L+L
Sbjct: 386 GKQREFQPYIEANWINNSKVYAVKMNGQTVGREGARNLG--EVRTGVEAKVNNNLSLWGN 443

Query: 661 FNRQTG-KHHQAKQGALNLQWTF 682
      Q G K +   QG L +++++
Sbjct: 444 VGVQLGDKGYSDTQGM LGVKYSW 466

```

```

tr Q8XAY3 Putative ATP-binding component of a transport system and      466
      adhesin                                                                AA
      protein [z2196] [Escherichia coli 0157:H7]                          align

```

Score = 129 bits (323), Expect = 2e-28
 Identities = 86/321 (26%), Positives = 156/321 (47%), Gaps = 19/321 (5%)

```

Query: 371 NAMIWRFMHNRLKPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGKTAPV 430
      N++   +H+RL P   +L   + S               +W+R + GH
Sbjct: 156 NSLFSHRLHDLRGEPQYTDLSHSQDSASS-----MWMRHVGGHERSSAGDGQLNT 205

Query: 431 EGYRKGVLGGGEVFTWQ-NESNQLSIGLMGGQAEQRSTFHNPDNDLTTGNVKGFGAGVY 489
      + R   +QLGG++ W N ++ +G+M G A Q S   +   + G + G+ AG+Y
Sbjct: 206 QANRYVLQLGGDLAQWSSNAQDRWHLGVMMAGYANQHSNTQSNRVGYKSDGRISGYSAGLY 265

Query: 490 ATWHQLQDKQTGAYADSWMQYQFRHRLNTED-GTERFTSKGITASIEAGYNALLAEHFT 548
      ATW+Q   +TGAY DSW Y F + +++++   + + S+G+TAS+E GY
Sbjct: 266 ATWYQNDANKTGAYVDSWALYNWFDNSVSSDNRSADDYDSRGVTASVEGGYTFEAGTCSG 325

Query: 549 KKGNSLRVYLQPQAQLTYLGV-NGKFSDSENAHLNLLGSRQLQTRVGVO----AKAQFSL 603
      +G      Y+QPQAQ+T++GV +   +   +   G   +QTR+GV+   +   Q
Sbjct: 326 SEGTLNTWYVQPQAQITWMGVKDSHDHARKDGTRIETEGDGNVQTRLGVKTYLNSHHQRDD 385

Query: 604 YKNIAIEPFAAVNALYHNKPFGEVMDGERRVINNKTAI-ESQLGVAVKIKSHLTLQATFN 662
      K   +P+   N + ++K + V+M+G+   +   + E + GV K+ ++L+L
Sbjct: 386 GKQREFQPYIEANWINNSKVYAVKMNGQTVSRDGARNLGEVVRTGVEAKVNNNLSLWGNVG 445

Query: 663 RQTG-KHHQAKQGALNLQWTF 682
      Q G K +   QG L +++++
Sbjct: 446 VQLGDKGYSDTQGM LGVKYSW 466

```

```

tr Q7AE01 Putative ATP-binding component of a transport system and      466
      adhesin                                                                AA
      protein [ECs2116] [Escherichia coli 0157:H7]                          align

```

Score = 129 bits (323), Expect = 2e-28
 Identities = 86/321 (26%), Positives = 156/321 (47%), Gaps = 19/321 (5%)

```

Query: 371 NAMIWRFMHNRLKPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGKTAPV 430
      N++   +H+RL P   +L   + S               +W+R + GH
Sbjct: 156 NSLFSHRLHDLRGEPQYTDLSHSQDSASS-----MWMRHVGGHERSSAGDGQLNT 205

```


Query: 431 EGYRKGVLGGEVFTWQ-NESNQLSIGLMGGQAEQRSTFHNPDNDLTTGNVKGFGAGVY 489
 + R +QLGG++ W N ++ +G+M G A Q S + + G + G+ AG+Y
 Sbjct: 206 QANRYVLQLGGDLAQWSSNAQDRWHLGVMAGYANQHSNTQSNRVGYKSDGRISGYSAGLY 265

Query: 490 ATWHQLQDKQTGAYADSWMQYQRFRRHINTED-GTERFTSKGITASIEAGYNALLAEHFT 548
 ATW+Q +TGAY DSW Y F ++++++ + + S+G+TAS+E GY
 Sbjct: 266 ATWYQNDANKTGAYVDSWALYNWFDNSVSSDNRSADDYDSRGVTASVEGGYTFEAGTCSG 325

Query: 549 KKGNSLRVYLQPQAQLTYLGV-NGKFSSENAHVNLGSRQLQTRVGVQ----AKAQFSL 603
 +G Y+QPQAQ+T++GV + + + + G +QTR+GV+ + Q
 Sbjct: 326 SEGTLNTWYVQPQAQITWMGVKDSHARKDGTRIETEGDGNVQTRLGVKTYLNSHHQRDD 385

Query: 604 YKNIAIEPFAAVNALYHNKPFGVEMDGERRVINNKTAI-ESQLGVAVKIKSHLTLOATFN 662
 K +P+ N + ++K + V+M+G+ + + E + GV K+ ++L+L
 Sbjct: 386 GKQREFQPYIEANWINNSKVYAVKMNGQTVSRDGARNLGEVRTGVEAKVNNNLSLWGNVG 445

Query: 663 RQTG-KHHQAKQGALNLQWTF 682
 Q G K + QG L +++++
 Sbjct: 446 VQLGDKGYSDTQGM LGVKYSW 466

tr Q83LF4 Hypothetical protein SF1157 [SF1157] [Shigella flexneri] 773 AA align

Score = 128 bits (321), Expect = 4e-28

Identities = 139/563 (24%), Positives = 231/563 (40%), Gaps = 85/563 (15%)

Query: 174 LGPKIKNSHINSEILSVGNYTEWANQVIHH-----IENYVSFAAHLYSGLDPFHYI 224
 L I +ILS G+ AN++++H ++ VS AA ++ H
 Sbjct: 199 LESSINGQEATVDILS-GSSLRANEILYHKDETSNVTITDSEVSSAADVFNNIKGHILT 257

Query: 225 EVTDNSHVIGQT-ISLDEFR-----LENSLWEPRWDSNVGKLKT'TNADIRFN-----T 271
 NS + G IS D+ +NS W+ + DS V L N+ + +
 Sbjct: 258 VDATNSKITGSANISTDDNTHTYLSLSDNSTWDIKADSTVSNLTVDNSTVYISRADGRDV 317

Query: 272 KSESLLVKEDYAGG---ARFRFAYDPKEAKNTALIFEKNVTGTSDIIFENPIDD-LKSLD 327
 + L + E+Y G R D + ++ N +GT+ + N +L+
 Sbjct: 318 EPTRLTITENYVGNNGVHLRLTELDDDNSATDKVINGNTSGTTRVKVTNAGGSGAYTLN 377

Query: 328 GHQIIKVNGTAD---KHAFRLSGKHQ-----KGIYTLSLQQRPEGFLPKCKN 371
 G +II V G ++ K + +G ++ K Y + Q G +
 Sbjct: 378 GIEIISVEGESNGEFIKDSRIFAGAYEYSLTRGNTATNKNWYLTNFQATSGGETNSGGS 437

Query: 372 AMIWRFMHNRLKPPIPYCALRLNNKNSDIFDRTLPRKG-----LWLRVI 415
 + L+P L N+ R R G LWLR I
 Sbjct: 438 SAPTVAPTPVLRPEAGSYVANLAAANTLFVRLNDRAGEMRYIDPVTEQERSRLWLRQI 497

Query: 416 DGHSNQWVQGTAPVEGYRKGVLGGEVFTWQ-NESNQLSIGLMGGQAEQRSTFHNPDND 474
 GH+ +R QLGG++ T +S+ +G+M G A + H+ +D
 Sbjct: 498 GGHNAWRDSNGQLRTTSHRYVSQLGDDLTTGGFTDSDSWRLGVMAGYARDYNLTHSSVSD 557

Query: 475 NLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFRRHINTED-GTERFTSKGITA 533
 + G+V+G+ AG+YATW + GAY DSW QY F++ + ++ E +++KG T
 Sbjct: 558 YRSKGSVRGYSAGLYATWFADDISKKGAYIDSWAQYSWFKNSVKGDELAYESYSAKGATV 617

Query: 534 SIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGKFSSENAHVNLGSR----- 587

```

      S+EAGY   L + F +           QPQAAQ  ++GV+           NAH   GSR
Sbjct: 618 SLEAGYGFALNKSFGLEAAKYTWIFQPQAAQAIWMGVD-----HNAHTEANGSRIENDAN 671

Query: 588 -QLQTRVG----VQAKAQFSLYKNIAIEPFPAAVNALYHNKPFQGVEMDG---ERRVINN-- 637
      +QTR+G    ++ + + S           EPF  +N ++++K F V M+G   E+  ++N
Sbjct: 672 NNIQTRLGFRTFIRTQEKNSGPHGDDFEPFVEMNWIHNSKDFAVSMNGVKVEQDGVSNLG 731

Query: 638 --KTAIESQLGVAVKIKSHLTLQ 658
      K +    L A +  ++ +Q
Sbjct: 732 EIKLGVNGNLPAAASVWGNVGVQ 754

```

```

tr   Q7UCU4           Hypothetical protein [S1242] [Shigella flexneri] 773 AA
                                align

```

Score = 128 bits (321), Expect = 4e-28

Identities = 139/563 (24%), Positives = 231/563 (40%), Gaps = 85/563 (15%)

```

Query: 174 LGPKIKNSHINSEILSVGNYTEWANQVIHH-----IENYVSFAAHLYSGLDPFHYI 224
      L   I           +ILS G+   AN++++H           ++ VS AA ++   H
Sbjct: 199 LESSINGQEATVDILS-GSSLRANEILYHKDETSNVTITDSEVSSAADVFINNKGHLT 257

Query: 225 EVTDNSHVIGQT-ISLDEFR-----LENSLWEPRWDSNVGKLKTTNADIRFN-----T 271
      NS + G   IS D+           +NS W+ + DS V L   N+ + +
Sbjct: 258 VDATNSKITGSANISTDDNTHTYLSLSDNSTWDIKADSTVSNLTVDNSTVYISRADGRDV 317

Query: 272 KSESLLVKEDYAGG---ARFRFAYDPKEAKNTALIFEKNVTGTSDIIFENPIDD-LKSLD 327
      +   L + E+Y G           R   D +   ++   N +GT+ +   N           +L+
Sbjct: 318 EPTRLTITENYVGNGVLHLRTELDDDNSATDKVINGNTSGTTRVKVTNAGGSGAYTLN 377

Query: 328 GHQIIKVN GTAD----KHAFLRSGKHQ-----KGIYTLSLQQRPEGFLPKCKN 371
      G +II V G ++   K +   +G ++           K Y + Q   G   +
Sbjct: 378 GIEIISVEGESNGEFIKDSRIFAGAYEYSLTRGNTTEATNKNWYLTNFQATSGGETNSGGS 437

Query: 372 AMIWRFMHNRLKPPPIPYCALRLNKNKNSDIFDRTLPRKG-----LWLRVI 415
      +           L+P           L   N+   R   R G           LWLR I
Sbjct: 438 SAPTVAPTPVLRPEAGSYVANLAAANTLFVMRLNDRAGETRYIDPVTEQERSRLWLRQI 497

Query: 416 DGHSNQWVQKKTAPVEGYRKGVLGGGEVFTWQ-NESNQLSIGLMGGQAEQRSTFHNPD TD 474
      GH+           +R   QLGG++ T   +S+   +G+M G A   + H+   +D
Sbjct: 498 GGHNAWRDSNGQLRTTSHRYVSQLGDDLTTGGFTDSDSWRLGVMAGYARDYNLTHSSVSD 557

Query: 475 NLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQFRHRINTED-GTERFTSKGITA 533
      + G+V+G+ AG+YATW           + GAY DSW QY F++ +   ++   E +++KG T
Sbjct: 558 YRSKGSVRGYSAGLYATWFADDISKKGAYIDSWAQYSWFKNSVKGDELAYESYSKATV 617

Query: 534 SIEAGYNALLAEHFTTKGNSLRVYLQPQAQLTYLGVNGKFSDSENAHVNLLGSR----- 587
      S+EAGY   L + F +           QPQAAQ  ++GV+           NAH   GSR
Sbjct: 618 SLEAGYGFALNKSFGLEAAKYTWIFQPQAAQAIWMGVD-----HNAHTEANGSRIENDAN 671

Query: 588 -QLQTRVG----VQAKAQFSLYKNIAIEPFPAAVNALYHNKPFQGVEMDG---ERRVINN-- 637
      +QTR+G    ++ + + S           EPF  +N ++++K F V M+G   E+  ++N
Sbjct: 672 NNIQTRLGFRTFIRTQEKNSGPHGDDFEPFVEMNWIHNSKDFAVSMNGVKVEQDGVSNLG 731

Query: 638 --KTAIESQLGVAVKIKSHLTLQ 658
      K +    L A +  ++ +Q

```

Sbjct: 732 EIKLGVNGNLPAAASVWGNVGVQ 754

tr Q9CKA8 Hypothetical protein PM1717 [PM1717] [Pasteurella multocida] 850 AA

align

Score = 126 bits (317), Expect = 1e-27

Identities = 117/434 (26%), Positives = 199/434 (44%), Gaps = 47/434 (10%)

Query: 274 ESLLVKEDYAG-----GARFRFAYDPKEAKNTALIFEKNVTGTSDIIFEN-PIDDLKS 325

+++ + EDY G A F +P + L N TG + + + D +

Sbjct: 429 KNV TINEDYQGHNGTLHLSADFN GTTNPTDT---LFIRGNATGKTRVAIHIGADAENA 484

Query: 326 LDGHQIIKVN GTADKHAFRLSGKHQKGIYTL SLQQRPE-----GFLPKCKNAM 373

++G +II+ N + D +AF + KG + L++R E G P + M

Sbjct: 485 VNGVKIIE T NTSTD-NAFVIDNYLSKGAFVYHLEKRHETN QDNWYLT SYIGGTPSYRAEM 543

Query: 374 IWRFMHNRLKPPIPYCALRLNNKNS-DIFDRTLPRKGLWLRVIDGHSNQWVQGKTAPVEG 432

N L LRL ++ S F K W+R ++G ++ ++ +

Sbjct: 544 A--SYANNLYAAHQLFQLRLEDRLSRHHFLNQ SADKTFWIRAVEGTNHNMRDNQNTTKA 601

Query: 433 YRKG VQLGGEVFTWQNESNQLSIGLMGGQAEQRSTFHNPD TDNLTTGNVKGFGAGVYATW 492

R QLG V N+++ + G+M G A+Q S + + G V+G+ G VY TW

Sbjct: 602 QRYVTQLGKTVI---NQAHYHA-GVMFGYAKQSSKTRSSRVGT-SRGKVQGYALGVYGTW 656

Query: 493 HQLQDKQTGAYADSWMQYQRFRRH-INTEDGTERFTSKGITASIEAGYNALLAEHFTKKG 551

+Q + TG Y DSW+QYQ F+++ IN ++ + ++G++AS+E GY+ L + +T

Sbjct: 657 YQNPNDTGLYIDSWLQYQWFKNQVINPASSSDNYRTQGLSASLELGYHLPLVQ-YTVAD 715

Query: 552 NSLRVYLQPQAQLTYLGVNGK-FSDSENAHVNLGSRQLQTRVG VQAKAQFSLYKNIA-- 608

+ +QPQAQ + +N K D + ++ +G + QTR+GV +FSL +

Sbjct: 716 LKHSLNIQPQAQFIWQKLNSKQHRDPQOTLIHYIGQONTQTRLGV----RFSLD SHFLNT 771

Query: 609 ---IEPFAAVNALYHNKPF GVEM-DGERRVINNKTAIESQLGVAVKIKSHLTLQATFN RQ 664

++P+ VN L+H K +G+ + D + K E + G+A + HL Q

Sbjct: 772 QWNLPKYFEVNW LHHAKDYGITINDV VNHIEGAKQLFEYKAGIASQFGRHLRFWLDTHQ 831

Query: 665 TGKHHQAKQGALNL 678

GK Q K LN+

Sbjct: 832 RGK-QQFKDNQLNV 844

tr P75997 Putative part of putative ATP-binding component of a transport system [b1170] [Escherichia coli K12] 338 AA

align

Score = 118 bits (296), Expect = 3e-25

Identities = 76/234 (32%), Positives = 119/234 (50%), Gaps = 19/234 (8%)

Query: 410 LWLRVIDGHSNQWVQGKTAPVEGYRKG VQLGGEVFTWQ-NESNQLSIGLMGGQAEQRSTF 468

LWLR I GH+ +R QLGG++ T +S+ +G+M G A +

Sbjct: 66 LWLRQIGGHNAWRDSNGQLRTTSHRYVSQLGGDLLTGGFTDSDSWRLGVMAGYARDYNLT 125

Query: 469 HNPDTDNLT TGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFRRHINTED-GTERFT 527

```

      H+  +D  + G+V+G+ AG+YATW      + GAY DSW QY  F++ +  ++  E ++
Sbjct: 126 HSSVSDYRSKGSVRGYSAGLYATWFADDISKKGAYIDSWAQYSWFKNSVKGDELAYESYS 185

Query: 528 SKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGKFSDSENAHVNLGSR 587
      +KG T S+EAGY  L + F  +      QPQAQ  ++GV+      NAH  GSR
Sbjct: 186 AKGATVSLEAGYGFAFNKSFGLAAKYTWIFQPQAQAIWMGVD-----HNAHTEANGSR 239

Query: 588 -----QLQTRVG-----VQAKAQFSLYKNIAIEPFAAVNALYHNKPFGEVMDG 630
      +QTR+G      ++ + + S      EPF  +N ++++K F V M+G
Sbjct: 240 IENDANNNIQTRLGFRTFIRTQEKNSGPHGDDFEPFVEMNWIHNSKDFAVSMNG 293

```

```

sp      Q03155      Adhesin aidA-I precursor [aidA-I] [Escherichia 1286 AA
      AIDA_ECOLI      coli]      align

```

Score = 118 bits (295), Expect = 4e-25

Identities = 105/414 (25%), Positives = 186/414 (44%), Gaps = 40/414 (9%)

```

Query: 302 LIFEKNVTGTSDIIFENPIDDL-KSLDGHQIIKVNGTADKHAFLSGKHQKGIYTLSQLQ 360
      L+ + N +G SDI++ N      ++ DG  II V G +D  F L  +  G Y  +LQ+
Sbjct: 880 LVVKGNTSGQSDIVYVNEGSGGQTRDGINIISVEGNSDAE-FSLKNRVVAGAYDYTLQK 938

Query: 361 RPEGFLPKCKNAMIWRFMH-----NRLKPPIPYCALRLNNKNS-----DIF 401
      E      N  +  H      + +P  A  +  NS      +
Sbjct: 939 GNES---GTDNKGWYLTSHLPTSDRQYRPENGSYATNMALANSLFLMDLNERKQFRAMS 995

Query: 402 DRTLPRKG-LWLRVIDGHSNQWV---QGKTAPVEGYRKGVLGGEVFTWQNES-NQLSIG 456
      D T P  +W+++ G S+  +  Q KT  +      QLGG+++ +  E  ++G
Sbjct: 996 DNTQPESASVWMKITGGISSGKLNDGQNKTTTNQFIN---QLGGDIYKFHAEQLGDFTLG 1052

Query: 457 LMGGQAEQRSTFHNPDTDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQFRHR 516
      +MGG A  +      N  ++      + G+  GVV TW+Q  +  TG +A++WMQY  F
Sbjct: 1053 IMGGYANAKGKTINYTSNKAARNTLDGYSVGVYGTWYQNGENATGLFAETWMQYNWFNAS 1112

Query: 517 INTED-GTERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGKFSD 575
      +  +  E++  G+TAS  GYN  +  + +G  +  +LQP  Q  ++GV
Sbjct: 1113 VKGDGLEEEKYNLNLGTASAGGGYNLNVHTWTSPEGITGEFWLQPHLQAVWMGVTPDTHQ 1172

Query: 576 SENAH-VNLLGSRQLQTRVGVQA--KAQFSLYKNIA--IEPFAAVNALYHNKPFGEVMDG 630
      +N  V  G  +QT+ G++A  K  + +L K+      P+  N  +++  FGV+M
Sbjct: 1173 EDNGTVVQGAGKNNIQTKAGIRASWKVKSTLDDKTGRRFRPYIEANWIHNTHFEGVKMSD 1232

Query: 631 ERRVIN-NKTAIESQLGVAVKIKSHLTLQATFNRTGKH-HQAKQGALNLQWTF 682
      + ++++ ++  E + G+  I  +L++      Q G H  A  GAL  ++++F
Sbjct: 1233 DSQLLSGSRNQGEIKTGIEGVITQNLVNGGVAYQAGGHGSNAISGALGIKYSF 1286

```

```

tr      Q7CPG9      Putative autotransported protein [misL] [Salmonella 955
      typhimurium]      AA
                        align

```

Score = 110 bits (276), Expect = 7e-23

Identities = 81/286 (28%), Positives = 142/286 (49%), Gaps = 22/286 (7%)

```

Query: 410 LWLRVIDGHS--NQWVQGKTAPVEGYRKGVLGGEVFTWQNES-NQLSIGLMGGQAEQRS 466

```

LW+R + H+ N + Y +QLGG++ W + ++ IG M G A ++
 Sbjct: 679 LWMRNVGAHTRFNDGSGQLKTRINSYV--LQLGGDLAQWSTDGLDRWHIGAMAGYANSQN 736
 Query: 467 TFHNPDTDNLTGTVKGFAGVYATWHQLQDKQTGAYADSWMQYQRFHRINTED-GTER 525
 + +D + G V G+ G+Y TW+ ++GAY D+WM Y F +++ +D E+
 Sbjct: 737 RTLSSVSDYHSRGQVTGYSVGLYGTWYANNIDRSAYVDTWMLYNWFDNKVMGQDQAAEK 796
 Query: 526 FTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGKFSDSENAHVNLL- 584
 + SKGITAS+EAGY+ L E + +LQP+AQ+ ++GV D+ A+ L+
 Sbjct: 797 YKSKGITASVEAGYSFRLGESAHQS-----YWLQPKAQVVMGVQA--DDNREANGTLVK 849
 Query: 585 --GSRQLQTRVGVOAKAQ----FSLYKNIAIEPFAAVNALYHNKPFVEMDG-ERRVINN 637
 + L TR+GV+A K+ +PF N +++ +P V+MD +
 Sbjct: 850 DDTAGNLLTRMGVKAYINGHNAIDNDKSREFQPFVEANWIHNTQPASVKMDDVSSDMRG 909
 Query: 638 KTAIESQLGVAVKIKSHLTLQATFNRTG-KHHQAKQGALNLQWTF 682
 K E ++G+ ++ L + +Q G + QG L ++++F
 Sbjct: 910 KNIGELKVGIEGQVTPRLNVWGNVAQQVGDGTGYSDTQGMGLMKYSF 955

tr Q9Z625 MisL [misL] [Salmonella typhimurium] 955 AA
 align

Score = 110 bits (276), Expect = 7e-23

Identities = 81/286 (28%), Positives = 142/286 (49%), Gaps = 22/286 (7%)

Query: 410 LWLRVIDGHS--NQWVQKTAPVEGYRKGVLGGEVFTWQNES-NQLSIGLMGGQAEQRS 466
 LW+R + H+ N + Y +QLGG++ W + ++ IG M G A ++
 Sbjct: 679 LWMRNVGAHTRFNDGSGQLKTRINSYV--LQLGGDLAQWSTDGLDRWHIGAMAGYANSQN 736
 Query: 467 TFHNPDTDNLTGTVKGFAGVYATWHQLQDKQTGAYADSWMQYQRFHRINTED-GTER 525
 + +D + G V G+ G+Y TW+ ++GAY D+WM Y F +++ +D E+
 Sbjct: 737 RTLSSVSDYHSRGQVTGYSVGLYGTWYANNIDRSAYVDTWMLYNWFDNKVMGQDQAAEK 796
 Query: 526 FTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGKFSDSENAHVNLL- 584
 + SKGITAS+EAGY+ L E + +LQP+AQ+ ++GV D+ A+ L+
 Sbjct: 797 YKSKGITASVEAGYSFRLGESAHQS-----YWLQPKAQVVMGVQA--DDNREANGTLVK 849
 Query: 585 --GSRQLQTRVGVOAKAQ----FSLYKNIAIEPFAAVNALYHNKPFVEMDG-ERRVINN 637
 + L TR+GV+A K+ +PF N +++ +P V+MD +
 Sbjct: 850 DDTAGNLLTRMGVKAYINGHNAIDNDKSREFQPFVEANWIHNTQPASVKMDDVSSDMRG 909
 Query: 638 KTAIESQLGVAVKIKSHLTLQATFNRTG-KHHQAKQGALNLQWTF 682
 K E ++G+ ++ L + +Q G + QG L ++++F
 Sbjct: 910 KNIGELKVGIEGQVTPRLNVWGNVAQQVGDGTGYSDTQGMGLMKYSF 955

sp P45508 Hypothetical protein yfaL precursor [yfaL] [Escherichia 1250
 YFAL_ECOLI coli] AA
 align

Score = 96.3 bits (238), Expect = 2e-18

Identities = 74/279 (26%), Positives = 136/279 (48%), Gaps = 27/279 (9%)

Query: 412 LRVIDGHSNQWVQKTAPVEGYRKGVLGGEVFT--WQNESNQLSIGLMGGQAEQRSTFH 469

```

      LRVI G +      G+ A E      VQL G++F+ W + + +G++GG ++ +
Sbjct: 989 LRVIGGDYHYTAAGQLAQHED-TSTVQLSGDLFSGRWGTDGEWM-LGIVGGYSNDNQGDSR 1046

Query: 470 NPDTDNLTGNGVKGFGAGVYATWHQLQDKQGTGAYADSWMQYQRFRRHRIN-TEDGTERFTS 528
      + T      G+ G+ ++W Q +++ GA+ DSW+QY F + ++ EDGT+ + S
Sbjct: 1047 SNMTGTRADNQNHGYAVGLTSSWFQHGNGKQKQGAWLDSWLQYAWFSNDVSEQEDGTDHYHS 1106

Query: 529 KGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGV-NGKFSDSENAHVNLGSR 587
      GI AS+EAGY L      V ++PQAQ+ Y GV      F+ + A V+
Sbjct: 1107 SGIIASLEAGYQWLPGR-----GVVIEPQAQVIYQGVQDDFTAANRARSQSQGD 1157

Query: 588 QLQTRVG VQAKAQFSLYKNIAIEPFAAVNALYHNKPFGVEMDGERRVINN---KTAIESQ 644
      +QTR+G+ ++ +      A+      ++ Y++ P      E++ +      I++ K      E +
Sbjct: 1158 DIQTRLGLHSEWR-----TAVHVIPTLDLNYHDPHSTEIEEDGSTISDDAVKQRGEIK 1211

Query: 645 LGVAVKIKSHLTLQATFNRTGKHHQAKQG---ALNLQW 680
      +GV I      ++L+ +      Q G      A+      ++ ++W
Sbjct: 1212 VGVGTGNISQVRSLRGSVAWQKGSDDFAQTAGFLSMTVKW 1250

```

tr Q8ZN57 Similar to the C-terminal region of AIDA [shdA] [Salmonella 2039 AA typhimurium]

align

Score = 95.5 bits (236), Expect = 3e-18

Identities = 75/289 (25%), Positives = 128/289 (43%), Gaps = 29/289 (10%)

```

Query: 410 LWLRVIDGHSNQWVQKTAPEGYRKGVLGGVFTWQNESNQLSIGLMGGQAEQRSTFH 469
      +W R      G +      ++      QLGG++ W N      +++G+M      +++
Sbjct: 1764 VWARFKAGKAESEAVSGNIDMDSNYSQFQLGGDILAWGNGQQSVTVGVM-----ASYI 1816

Query: 470 NPDTDNL-----TTGNGVKGFGAGVYATWHQLQDKQGTGAYADSWMQYQRFRRHRIN 518
      N DTD+      ++GNV G+      G VYATW      +GAY DSW QY + + +
Sbjct: 1817 NADTDSTGNRGADGSQFTSSGNVDGYNLGVYATWFADAQTHSGAYVDSWYQYGFYNNNSVE 1876

Query: 519 TED-GTERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNG-KFSDS 576
      + D G+E + S      S+E GY      +A      GN+ V L PQAQ+ + +      D+
Sbjct: 1877 SGDAGSESYDSTANAVSLETGYRYDIA---LSNGNT--VSLTPQAQVWQNYSDSVKDN 1931

Query: 577 ENAHVNLLGSRQLQTRVG VQAKAQFSLYKNIAIEPFAAVNALYHNKPFGVEMDGERRVIN 636
      ++      TR+G++ +      I+PFA N L+ +      V D + V
Sbjct: 1932 YGTRIDGQDGDSWTTTLGLRVDGKLYKGSRTVIQPFANLHTSDDVSVSFD-DATVKQ 1990

Query: 637 NKTA--IESQLGVAVKIKSHLTLQATFNRTGKHHQAK-QGALNLQWTF 682
      + A      E ++G+ I      +++A      QTG +      G+LNL++ +
Sbjct: 1991 DLPANRAELKVGLQADIDKQWSVRAQVAGQTGSNDFGDLNGSLNLRYNW 2039

```

tr Q8CVV7 Hypothetical protein yfaL [yfaL] [Escherichia coli 1254 AA 06]

align

Score = 95.5 bits (236), Expect = 3e-18

Identities = 71/278 (25%), Positives = 138/278 (49%), Gaps = 25/278 (8%)

```

Query: 412 LRVIDGHSNQWVQKTAPEGYRKGVLGGVFTWQ-NESNQLSIGLMGGQAEQRSTFHN 470

```

```

          LRVI G +      G+ A E      VQL G +F+      + + +G++GG ++ +      +
Sbjct: 993 LRVIGGRYHYTAVGQLAQHED-TSTVQLSGNLFSGHWGDDGEWMLGIVGGYSDNQGDSRS 1051

Query: 471 PDTDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFRRHRIN-TEDGTERFTSK 529
          T          G+  G+ ++W+Q  +++ GA+ DSW+QY  F + ++ +DG + + S
Sbjct: 1052 NMTGTRADNQNHGYAVGLTSSWYQHGNQKQGAWLDSWLQYAWFNNDVSEQDDGVDHYHSS 1111

Query: 530 GITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGV-NGKFSDSENAHVNLLGSRQ 588
          GI AS+EAGY  L          V ++PQAQ+ Y GV      F+ + +A V+
Sbjct: 1112 GIIASLEAGYQWLPGR-----GVVIEPQAQVIYQGVQDDFTAANHARVSQSQGDD 1162

Query: 589 LQTRVGQAKAQFSLYKNIAIEPFAAVNALYHNKPFGEVMDGERRVINNKTAI---ESQL 645
          +QTR+G+ ++ + +      +++ P  +N  Y++ P  E++ +  I++  A  E ++
Sbjct: 1163 IQTRLGLHSEWRTA----VSVTPTLDLN--YYHDPHATEIEEDGSTISDDAAKQRGEIKV 1216

Query: 646 GVAVKIKSHLTLQATFNRQTGKHHQAKQG---ALNLQW 680
          G+  I  ++L+ +  Q G  A+      ++ ++W
Sbjct: 1217 GITGNISQRVSLRGSVAWQKGSDDFAQTAGFLSMTVKW 1254

```

```

tr    Q9XCJ4          ShdA [shdA] [Salmonella typhimurium]          2035 AA
                                     align

```

Score = 95.5 bits (236), Expect = 3e-18

Identities = 75/289 (25%), Positives = 128/289 (43%), Gaps = 29/289 (10%)

```

Query: 410 LWLRVIDGHSNQWVQKGTAPVEGYRKGVLGGGEVFTWQNESNQLSIGLMGGQAEQRSTFH 469
          +W R  G +          ++      QLGG++  W N      +++G+M      +++
Sbjct: 1760 VWARFKAGKAESEAVSGNIDMSNYSQFQLGGDILAWGNGQQSVTVGVM-----ASYI 1812

Query: 470 NPDTDNL-----TTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFRRHRIN 518
          N DTD+          ++GNV G+  GVIATW          +GAY DSW QY  + + +
Sbjct: 1813 NADTDSTGNRGADGSQFTSSGNVDGYNLGVYATWFADAQTHSGAYVDSWYQYGFYNNNSVE 1872

Query: 519 TED-GTERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNG-KFSDS 576
          + D G+E + S      S+E GY  +A      GN+  V L PQAQ+ +  +  D+
Sbjct: 1873 SGDAGESYDSTANAVSLETGYRYDIA--LSNGNT--VSLTPQAQVWQNYADSVDKDN 1927

Query: 577 ENAHVNLLGSRQLQTRVGQAKAQFSLYKNIAIEPFAAVNALYHNKPFGEVMDGERRVIN 636
          ++          TR+G++  +          I+PFA  N L+ +  V  D +  V
Sbjct: 1928 YGTRIDGQDGDSWTTTRLGLRVDGKLYKGSRTVIQPF AEANWLHTSDDVSVSFD-DATVKQ 1986

Query: 637 NKTA--IESQLGVAVKIKSHLTLQATFNRQTGKHHQAK-QGALNLQWTF 682
          +  A  E ++G+  I      +++A  QTG +      G+LNL++ +
Sbjct: 1987 DLPANRAELKVGLQADIDKQWSVRAQVAGQTGSNDFGDLNGSLNLRYNW 2035

```

```

tr    Q8XE28 Putative ATP-binding component of a transport system [yfaL] 1250 AA
          [Escherichia coli O157:H7]
                                     align

```

Score = 93.2 bits (230), Expect = 1e-17

Identities = 72/279 (25%), Positives = 134/279 (47%), Gaps = 27/279 (9%)

```

Query: 412 LRVIDGHSNQWVQKGTAPVEGYRKGVLGGGEVFT--WQNESNQLSIGLMGGQAEQRSTFH 469

```

```

          LRVI G +   G+ A E   VQL G++F+ W +   + +G++GG ++ +
Sbjct: 989  LRVIGGDYHYTAAGQLAQHED-TSTVQLSGDLFSGRWGTDGEWM-LGIVGGYSDNQGDSR 1046

Query: 470  NPDTDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFHRINT-EDGTERFTS 528
          + T           G+ G+ ++W Q   ++ GA+ D+W+QY F + ++ EDG + + S
Sbjct: 1047 SSMTGTRADNQNHHGYAVGLTSSWFQHGKQKQGAWLDNWLQYAWFSNDVSEHEDGVDHYHS 1106

Query: 529  KGITASIEAGYNALLAEHFTTKGNSLRVYLQPQAQLTYLGV-NGKFSDSENAHVNLLGSR 587
          GI AS+EAGY L           V ++PQAQ+ Y GV   F+ + A V+
Sbjct: 1107 SGIIASLEAGYQWLPGR-----GVVIEPQAQVIYQGVQDDFTAANRARVSQSQGD 1157

Query: 588  QLQTRVGQVQAKAQFSLYKNIAIEPFAAVNALYHNKPFVEMDGERRVINN---KTAIESQ 644
          +QTR+G+ ++ +   A+   ++ Y++ P   E++ +   I++ K E +
Sbjct: 1158 DIQTRLGLHSEWR-----TAVHVIPTLDLNYHDPHSTEIEEDASTISDDAVKQGEIK 1211

Query: 645  LGVAVKIKSHLTLQATFNROTGKHHQAKQG---ALNLQW 680
          +GV I   ++L+ +   Q G   A+   ++ ++W
Sbjct: 1212 VGVGTGNISQVRSLRGSVAWQKGSDDFAQTAGFLSMTVKW 1250

```

```

tr  Q7AC44  Putative ATP-binding component of a transport system      1250
          [ECs3116]
          [Escherichia coli O157:H7]
                                     AA
                                     align

```

Score = 93.2 bits (230), Expect = 1e-17

Identities = 72/279 (25%), Positives = 134/279 (47%), Gaps = 27/279 (9%)

```

Query: 412  LRVIDGHSNQWVQKKTAPVEGYRKGVLGGEVFT--WQNESNQLSIGLMGGQAEQRSTFH 469
          LRVI G +   G+ A E   VQL G++F+ W +   + +G++GG ++ +
Sbjct: 989  LRVIGGDYHYTAAGQLAQHED-TSTVQLSGDLFSGRWGTDGEWM-LGIVGGYSDNQGDSR 1046

Query: 470  NPDTDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFHRINT-EDGTERFTS 528
          + T           G+ G+ ++W Q   ++ GA+ D+W+QY F + ++ EDG + + S
Sbjct: 1047 SSMTGTRADNQNHHGYAVGLTSSWFQHGKQKQGAWLDNWLQYAWFSNDVSEHEDGVDHYHS 1106

Query: 529  KGITASIEAGYNALLAEHFTTKGNSLRVYLQPQAQLTYLGV-NGKFSDSENAHVNLLGSR 587
          GI AS+EAGY L           V ++PQAQ+ Y GV   F+ + A V+
Sbjct: 1107 SGIIASLEAGYQWLPGR-----GVVIEPQAQVIYQGVQDDFTAANRARVSQSQGD 1157

Query: 588  QLQTRVGQVQAKAQFSLYKNIAIEPFAAVNALYHNKPFVEMDGERRVINN---KTAIESQ 644
          +QTR+G+ ++ +   A+   ++ Y++ P   E++ +   I++ K E +
Sbjct: 1158 DIQTRLGLHSEWR-----TAVHVIPTLDLNYHDPHSTEIEEDASTISDDAVKQGEIK 1211

Query: 645  LGVAVKIKSHLTLQATFNROTGKHHQAKQG---ALNLQW 680
          +GV I   ++L+ +   Q G   A+   ++ ++W
Sbjct: 1212 VGVGTGNISQVRSLRGSVAWQKGSDDFAQTAGFLSMTVKW 1250

```

```

tr  Q9FCW0  Hypothetical protein [Kluyvera ascorbata]      652 AA
                                     align

```

Score = 73.9 bits (180), Expect = 9e-12

Identities = 76/299 (25%), Positives = 123/299 (40%), Gaps = 44/299 (14%)

```

Query: 246  NSLWEPRWDSNVGKLKTTNADIRFNTKSES--LLVKEDYAGG---ARFRFAYDPKEAKNT 300

```



```

      NSLW+  DS +  L  N  +  +  S  L  V  +Y  G  R  +
Sbjct: 361 NSLWKMTGDSVLSSLTLNNGTVEWAGASAGNVLTVAGNYQGNNGLLRINTVLGDDNSVTD 420

Query: 301 ALIFEKNVTGTSDIIFENPIDDL-KSLDGHQIIKVNGTADKHAFLRLSGKHQKGIYTLSLQ 359
      L+ E + +GT+ +  N      K+++G ++I V+G +D + F  SG+  G Y  +L+
Sbjct: 421 KLVVEGDTSGTTFVAVTNAGGSGDKTINGIEVIHVDGASDGN-FVQSGRIVAGSYEYALR 479

Query: 360 QRPEGFLPKCKNAMIWRFMHNRLKP--PIP-----YCALRLNNKNSDIFDRTLTP 406
      R EG      W M+  P  P+P      Y A L  N+  R
Sbjct: 480 -RGEGTAFNH-----WYLMNAATTPEEPVPEKPTLRPESGSYIA-NLAAANTMFTTRLHD 532

Query: 407 RKG-----LWLRVIDGHSNQWVQGKTAPVEGYRKGVLGGGEVFTWQ-NE 449
      R G      +W+R + GH+      + R  +QLGG++ W N
Sbjct: 533 RLGETQYIARLTGEQKVTSMWMNRNVGGHTRFKDSSGQLSTQSNRYVLQLGGDIAQWSTNG 592

Query: 450 SNQLSIGLMGGQAEQRSTFHNPDNDLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWM 508
      ++ +GLM G  +S  +  +  + G V G+ AG+Y TW+  +  +TG+Y DSW+
Sbjct: 593 LDRWHLGLMAGYGNQSNTKSDVSRYSRGOVTGYSAGIYGTWYANEADKTGSYLD SWI 651

```

tr Q6IU25 **YdeU (Fragment) [Escherichia coli B]** 169 AA
align

Score = 70.5 bits (171), Expect = 1e-10

Identities = 39/138 (28%), Positives = 63/138 (45%), Gaps = 11/138 (7%)

```

Query: 371 NAMIWRFMHNRLKPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGKTAPV 430
      N++  +H+RL P      + +D      +W+R + GH
Sbjct: 41  NSLFSHRLHDLRLGEP-----QYTDLSHSQGSASSMWMRHVGGHERSRAGDQQLNT 90

Query: 431 EGYRKGVLGGGEVFTWQ-NESNQLSIGLMGGQAEQRSTFHNPDNDLTTGNVKGFGAGVY 489
      + R  +QLGG++ W N  ++  +G+M G A Q S  +      + G + G+ AG+Y
Sbjct: 91  QANRYVLQLGGDLAQWSSNAQDRWHLGVMAGYANQHSNTQSNRVGYKSDGRISGYSAGLY 150

Query: 490 ATWHQLQDKQTGAYADSW 507
      ATW+Q  +TGAY DSW
Sbjct: 151 ATWYQNDANKTGAYVDSW 168

```

tr Q88LP9 **Outer membrane autotransporter [PP1880] [Pseudomonas putida (strain KT2440)]** 730 AA
align

Score = 66.2 bits (160), Expect = 2e-09

Identities = 109/474 (22%), Positives = 183/474 (37%), Gaps = 84/474 (17%)

```

Query: 228 DNSHVIGQTIS-----LDEF-----RLENSL-----WEPRWDSNVGKLKTTN 264
      +NSH++G ++      L+ F      RL+N      W      +S V L  N
Sbjct: 257 NNSHLVGDIVAASGGTANVLLENFATLKGRLDNVASLEINSGGEWTLVDNSQVTDLSLDN 316

Query: 265 ADIRFNTKSESLLVK-EDYAGGARFRFAYDPKEAKNTALIFEKNVTGTSD-----IIFEN 318
      +RF  E  +  E+  G  F  D  +++  L  +VTGT+      +I  +
Sbjct: 317 GAVRFGGPGFEFFTLSENLTGNGTFFIMEADFSTSQSDFL----DVTGTASGNHQLLISAS 372

Query: 319 PIDDLKSLDGHQIIKVNGTADKHAFLRLSGKHQKGIYTLSLQQRPEG-----FLPK 368

```

```

      D L      H +      G D      L G      G Y+ L Q R +      P
Sbjct: 373 GNDPLTDNSLHVHTAAG--DSQFSLGGSVDLGAYSVDLVQRGDNDWYLDATTRTVSPG 430

Query: 369 CKNAMIWRFMHNRLKPPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVGK- 427
      + M      + P I Y L +      R + G W+R      + NQ+ T
Sbjct: 431 TQTVMALA----NVVPTIWIYGELGVLRSRMGDVRRNPGKAGGWVR---SYGNQFNVSATS 483

Query: 428 -APVEGYRKGVQLGGEVFTWQNESNQLSIGLMGGQAEQRSTFHNPDNDNLTTGN---VKG 483
      A + ++G+ +G +      + N L +G+ G      + N D NL G+ V
Sbjct: 484 GAAYQQQQQLSIGADAPLAAGDGNWL-VGITAG-----YSNSDL-NLARGSSASVDS 534

Query: 484 FGAGVYATWHQLQDKQTGAYADSWMQYQFRHR--INTEDGTER---FTSKGITASIEAG 538
      + AG YATW D ++G Y D+ + RFR++ + DG++ +++ G S+E G
Sbjct: 535 YHAGAYATW---LDPESGYIYDITVARINRFRNQADVRLSDGSKAKGDYSNLGAGVSLEVG 591

Query: 539 YNALLAEHFTKKGNLSRVYLQPQAQLTYLGVNGK-FSDSENAHVNLLGSRQLQTRVGVQA 597
      + LA+ +      +L+P AQL+ L V GK +S      N + L +VG
Sbjct: 592 RHLNLADDW-----FLEPFAQLSGLVVQGGDYSLDNGMRANSNSTHSLGKVGTSV 642

Query: 598 KAQFSLYKNIAIEPFAAVNALYHNKPFVEMDGERRVINNKTAIESQLGVAVKI 651
      FS      +++P+ V A++      + R +N      +++G V +
Sbjct: 643 GRTFSAGTGRSVQPYLRVAAVHEFVNDNQVKVNDNRFSSNLAGSRAEIGAGVAV 696

```

```

tr Q884S7 Autotransporter, putative [PSPTO2011] [Pseudomonas syringae] 769
      (pv. AA
      tomato)] align

```

Score = 66.2 bits (160), Expect = 2e-09

Identities = 105/460 (22%), Positives = 184/460 (39%), Gaps = 58/460 (12%)

```

Query: 249 WEPRWDSNVGKLKTTNADIRFNTKSESLLVKEDYAGGARFRFAYDPKEAKNTALIFEKNV 308
      W+ + D+ V L      + F + L + +G F D A +
Sbjct: 342 WQMQGDNAVTSLSMQGGSVGFGEFHTLSLNELSGSGTFGLRVDLDNAVGD LINVNGQA 401

Query: 309 TGTSDIIFENPIDDLKSLDGHQIIKVNGT--ADKHAFRLSGKHQKGIYTLSQLQRPEGFL 366
      +G + N ++ S D Q +KV T D L G+ G Y+ L+Q+ +
Sbjct: 402 SGQFGLRVRNTGVEVISAD-MQPLKVVHTEGGDAQFSLGGRVDLGAYSYLLEQQGNDWF 460

Query: 367 PKCKNAMIWRFMHNRLK----PPIPYCALR-LNNKNSDIFDRTLPRKGLWLRVIDGHSN 420
      ++ +I      + L      P I L L ++ ++ R + G W+R N
Sbjct: 461 VVGRDKVIS PSTQSALALYSAAPAIWMSELSTLRSRMGEV--RASGQAGGWMRAYGSRLN 518

Query: 421 QWVQGTAPVEGYRK--GVQLGGEVFTWQNESNQLSIGLMGGQAEQRSTFHNPDNDNL 477
      T+ YR+ G+ LG + + + +L +G++GG + + D T
Sbjct: 519 ----ATTSDGVDYRQKQSGLSLGADAPV-EVSNRLLVGVLGGYSTS-----DL DVS RGT 568

Query: 478 TGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQFRHR--INTEDGTER---FTSKGIT 532
      TG V + AG Y TW L D G Y D ++ RFR++ + D ++ +++ G+
Sbjct: 569 TGKVASYYAGAYGTW--LSD--DGYVVDGV LKLNFRNKADVAMS DASKAKGDYSNTGVG 624

Query: 533 ASIEAGYNALLAEHFTKKGNLSRVYLQPQAQLTYLGVNGKFSDSENAHVNLLGSRQLQTR 592
      +EAG + LA+ +      +L+P AQL+ + V G+ +N + ++ +TR
Sbjct: 625 GWVEAGRHIKLADDY-----FLEPFAQLSSVVVQGGELRLDNG----MKAKNARTR 671

Query: 593 VGVQAKAQFSLYKNIAIEPFAAVNALYHNKPFVEMDGERRVINNKTAIESQL----- 645

```

V K SL + +A++ V Y E V N ++ L
 Sbjct: 672 -SVLGKVGTSLGRTVALKD-GGVLQPYVRVAVAQEFSSRRNEVKANDVKFDNSLFGSRGEL 729
 Query: 646 --GVAVKIKSHLTLQATFNRTGKHHQAKQGA-LNLQWTF 682
 GV+V + L L A + G+H + GA + L+ TF
 Sbjct: 730 GAGVSVSLSERLKLHADVDYMKGQHIEQPWGANVGLRLTF 769

tr Q7VYJ2 Autotransporter [BP1344] [Bordetella pertussis] 866 AA
align

Score = 65.5 bits (158), Expect = 3e-09

Identities = 79/308 (25%), Positives = 131/308 (41%), Gaps = 46/308 (14%)

Query: 397 NSDIFDRTLPRKGLWLRVIDGHSNQWVQGTAP-VEGYRKGVLGGGEVFTW-QNESNQLS 454
 N + R R W R G+S Q + G P ++ GVQLG ++++ + + Q
 Sbjct: 583 NQALLARDGERVAAWARAYGCNSKQALDGDGAPGIDARLAGVQLGQDLYSSVRPDGGQHR 642
 Query: 455 IGLMGGQAEQRSTFHN--DTDNLTTGN--VKGFGAGVYATWHQLQDKQTGAY-ADSWMQ 509
 GL GG + R H + TG + G+ G Y T+ + A A++WM
 Sbjct: 643 FGLFGGYGQARGDTHGSAGGERDAATGRLTIDGYSVGGYWTYVGPRGWYVDAVLANTWMD 702
 Query: 510 YQFRHRINTEDGTERFT-SKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTY-- 566
 +++ G + T + TAS+E+GY L+E +T L+PQAQL Y
 Sbjct: 703 IDT-----DSKAGRDADTRGQAFTASLESGYPLALSERWT-----LEPQAQLIYQH 748
 Query: 567 LGVNGKFSDSENAHVNLLGSRQLQTRVGVAQAQFSLYKNIAIEPFAAVNALYHNKPFQV 626
 V+G FSD+ + V + L R+G + + +++ + P+AA+N F
 Sbjct: 749 TRVDG-FSDAV-SEVRIRDDNALTARLGARLQGEYAAAAQV-WRPYAALN-----FWR 798
 Query: 627 EMDGERRVI-----NNKTAIESQLGVAVKIKSHLTLQATFNRTGKHHQAKQGA-- 675
 GE V+ T++E G +V + L L T Q +GA
 Sbjct: 799 TFSGENTVVLGEDSIDTRRGATSLELAAGASVTLARSLALYGRLAYATSIDSQYLRGASA 858
 Query: 676 -LNLQWTF 682
 L +++T+
 Sbjct: 859 QLGMRYTW 866

tr Q7WIM1 Putative autotransporter [BB2830] [Bordetella bronchiseptica] 849 AA
align
(Alcaligenes bronchisepticus)]

Score = 64.3 bits (155), Expect = 7e-09

Identities = 75/304 (24%), Positives = 125/304 (40%), Gaps = 38/304 (12%)

Query: 397 NSDIFDRTLPRKGLWLRVIDGHSNQWVQGTAP-VEGYRKGVLGGGEVFTW-QNESNQLS 454
 N + R R W R G+S Q + G P ++ GVQLG ++++ + + Q
 Sbjct: 566 NQALLARDGERVAAWARAYGCNSKQALDGDGAPGIDARLAGVQLGQDLYSSVRPDGGQHR 625
 Query: 455 IGLMGGQAEQRSTFHN--DTDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQR 512
 GL GG R H + TG + G V W + + G Y D+ +
 Sbjct: 626 FGLFGGYGHARGDTHGSAGGERDAATGRLTIDGYSVGGYWTYVGPR--GWYVDTVLANW 683
 Query: 513 FRHRINTEDGTERFT-SKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVN- 570

```
      +++ G + T + +TAS+E+GY L+E +T L+PQAQL Y
Sbjct: 684 MDIDTDSKAGRDADTRGQALTASLESGYPLALSERWT-----LEPQAQLIYQHTRV 734

Query: 571 GKFSSENHVNLLGSRQLQTRVGVQAKAQFSLYKNIAIEPFAAVNALYHNKPFVEMDG 630
      FSD+ + V + L R+G + + +++ + P+AA+N F G
Sbjct: 735 DDFSADAV-SEVRIRDDNALTARLGARLQGEYAAAAQV-WRPYAALN-----FWRTFSG 785

Query: 631 ERRVI-----NNKTAIESQLGVAVKIKSHLTLQATFNRTGKHHQAKQGA---LNL 678
      E V+ T++E G +V + L L T Q +GA L +
Sbjct: 786 ENTVVLGEDSIDTRRGATSLELAAGASVTLARSLALYGRLAYATSIDSQYLRGASACLGM 845

Query: 679 QWTF 682
      ++T+
Sbjct: 846 RYTW 849
```

tr Q883X2 Autotransporter, putative [PSPT02225] [*Pseudomonas syringae* 773
(pv. AA
tomato)] align

Score = 63.9 bits (154), Expect = 9e-09
Identities = 69/283 (24%), Positives = 116/283 (40%), Gaps = 39/283 (13%)

```
Query: 407 RKGLWLRVIDGHSNQWVQGKTAPVEGYRK---GVQLGGEVFTWQNESNQLSIGLMGGQAE 463
      + G W+R N A GY++ G LG + + L+ G+M GQ+
Sbjct: 508 QSGAWMRTYGNKFNV---SDASGFGYQQTQQGFSLGADGKVPMDGQWLA-GVMACQSS 562

Query: 464 QRSTFHNPDNDLTGNNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQFRH--RINTED 521
      + D +G V + G Y+TW D TG Y D +++ RF + R+N D
Sbjct: 563 SDLSTL-----DRGASGKVDSYYVGAYSTW---LSDTGYDFDGVLFKFNRFNNKARVNLSD 614

Query: 522 GTER---FTSKGITASIEAGYNALLAEHFTKKGNLSLRVYLQPQAQLTYLGVNGKFSDSEN 578
      GT +++ G+ AS+E G + L + +++P +QL + V GK + +N
Sbjct: 615 GTRTKGDYSNSGVGASLEFGRHIKLDNGY-----FVEPYSQLAGVVVEGKDYELDN 665

Query: 579 ---AHVNLLGSRQLQTRVGVQAKAQFSLYKNIAIEPFAAVNALYHNKPFVEMDGERRVI 635
      A +L +R L ++G F L + ++P+ A H E+ V
Sbjct: 666 GMRAENDL--TRSLVGKLGATTGRNFDLQGRTVQPYVR-TAWVHEFAKNNEVQVNDNVF 722

Query: 636 NNK---TAIESQLGVAVKIKSHLTLQATFNRTGKHHQAKQGA 675
      NN + E +G+A + + A F G + GA
Sbjct: 723 NNDLSGSRGELGIGIAASLSERFQVHADFEHSNGDKVEQPWGA 765
```

tr Q7WIB0 Putative autotransporter [BB2941] [*Bordetella bronchiseptica* 937 AA
(*Alcaligenes bronchisepticus*)] align

Score = 55.5 bits (132), Expect = 3e-06
Identities = 106/493 (21%), Positives = 170/493 (33%), Gaps = 76/493 (15%)

```
Query: 224 IEVTDNSHVIGQTISLDEFRL-ENSLWEPRWDSNVGKLKTT-NADIRFNTKSE---SLLV 278
      ++V D + G T SLD L W S+VG+L A + F + L
Sbjct: 477 LDVADGAQWHGATQSLDRLALGAGGQWRMSAASSVGELSMEPGAADVFGDAAGPGFQTLT 536
```

Query: 279 KEDYAGGARFRFAYDPKEAKNTALIFEKNVTGTSDIIFENPIDDLKSLDGHQIIKVNHTA 338
AG F D L+ G + P S Q + V A
Sbjct: 537 VRTLAGAGSFEMRADAALHADQLVVTDAQEGRHRVWLRAPAGAEPKSA--QAVLVRAPA 594

Query: 339 D-KHAFRLSGKHQK---GIYTLSQLQRPEGFLPKCKNAM-----IWR 376
D K +F L G + G Y L Q+P G + +W
Sbjct: 595 DGKASFELDGS DGRADFGTYRYGLAQPPGAWGLVRTGYSSTAAAALDTGGLGAVQGLWY 654

Query: 377 FMHNRLKPPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGTAPVEGYRKG 436
N L + LRLN + R ++ R+ + QG + G
Sbjct: 655 AESNALGKRMG--ELRLNPDAGGAWGRAFSQRQ---RISPRAGRHFQQGVS-----G 701

Query: 437 VQLGGEVFTWQNESNQLSIGLMGGQAEQRSTFHNPDTDNLTGTVKGFAGVYATWHQLQ 496
++LG + W + G + G F + + +V G+ + A
Sbjct: 702 IELGADR-AWPVAGGRWHAGGLLGYTRASRGFSGQKGHTDSVHVGGYATYIGAN----- 755

Query: 497 DKQTGAYADSWMQYQFRHRINTEDGTER-----FTSKGITASIEAGYNALLAEHFTKKG 551
G YAD+ ++ RF + + R + + G+ ++EAG L H+
Sbjct: 756 ----GVYADATLRASRFENSFDVPGWAGRTVSGSYRANGVGVTLGARRLALDRHW---- 807

Query: 552 NSLRVYLQPPAQLTYLGVNG-KFSDSENAHVNLGSRQLQTRVGVAQAQFSLYKNIAIE 610
+++PQA+L + G ++ S + G LQ RVG QA +F L ++
Sbjct: 808 -----FVEPQAEALAWFRAGGGTYTASNGLRIEDDGGTSLQARVGAQAGRRFDLGGAVVQ 862

Query: 611 PFAAVNALYHNKPFPG-VEMDG-ERRVINNKTAIESQLGVAVKIKSHLTLQATFNRQTGKH 668
P+A ++ + K V +G R +E LGVA + L A++ G
Sbjct: 863 PYAQLSWVQELKGVSTVRTNGIAHRTDLGAGRVELGLGVAAALGKGHNLYASYEYAHGPR 922

Query: 669 HQAKQGALNLQWT 681
L+L WT
Sbjct: 923 -----LSLPWT 928

tr Q7W6E1 Putative autotransporter [BPP2975] [Bordetella parapertussis] 937 AA

align

Score = 55.5 bits (132), Expect = 3e-06

Identities = 106/493 (21%), Positives = 170/493 (33%), Gaps = 76/493 (15%)

Query: 224 IEVTDNSHVIGQTISLDEFRL-ENSLWEPRWDSNVGKLKTT-NADIRFNTKSE---SLLV 278
++V D + G T SLD L W S+VG+L A + F + L
Sbjct: 477 LDVADGAQWHGATQSLDRLALGAGGQWRMSAASSVGELSMEPGAHVFGDAAGPGFQTLT 536

Query: 279 KEDYAGGARFRFAYDPKEAKNTALIFEKNVTGTSDIIFENPIDDLKSLDGHQIIKVNHTA 338
AG F D L+ G + P S Q + V A
Sbjct: 537 VRTLAGAGSFEMRADAALHADQLVVTDAQEGRHRVWLRAPAGAEPKSA--QAVLVRAPA 594

Query: 339 D-KHAFRLSGKHQK---GIYTLSQLQRPEGFLPKCKNAM-----IWR 376
D K +F L G + G Y L Q+P G + +W
Sbjct: 595 DGKASFELDGS DGRADFGTYRYGLAQPPGAWGLVRTGYSSTAAAALDTGGLGAVQGLWY 654

Query: 377 FMHNRLKPPPIPYCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGTAPVEGYRKG 436
N L + LRLN + R ++ R+ + QG + G
Sbjct: 655 AESNALGKRMG--ELRLNPDAGGAWGRAFSQRQ---RISPRAGRHFQQGVS-----G 701

Query: 437 VQLGGEVFTWQNESNQLSIGLMGGQAEQRSTFHNPDNDLTTGNVKGFGAGVYATWHQLQ 496
++LG + W + G + G F + + +V G+ + A
Sbjct: 702 IELGADR-AWPVAGGRWHAGLLGYTRASRGFSGQGKHTDSVHVGGYATYIGAN----- 755

Query: 497 DKQTGAYADSWMQYQRFRRHINTEDGTER-----FTSKGITASIEAGYNALLAEHFTKKG 551
G YAD+ ++ RF + + R + + G+ ++EAG L H+
Sbjct: 756 ----GVYADATLRASRFENSFDVPGWAGRTVSGSYRANGVGVTLAAGRRLALDRHW---- 807

Query: 552 NSLRVYLQPQAQLTYLGVNG-KFSDSENAHVNLLGSRQLQTRVGVQAKAQFSLYKNIAIE 610
+++PQA+L + G ++ S + G LQ RVG QA +F L ++
Sbjct: 808 -----FVEPQAEALWFRAGGGTYTASNGLRIEDDGGTSLQARVGAQAGRRFDLRGGAVVQ 862

Query: 611 PFAAVNALYHNKPFG-VEMDG-ERRVINNKTAIESQLGVAVKIKSHLTLQATFNRQTGKH 668
P+A ++ + K V +G R +E LGVA + L A++ G
Sbjct: 863 PYAQLSWVQELKGVSTVRTNGIAHRTDLGAGRVELGLGVAAALGKGHNLYAS'YEYAHGPR 922

Query: 669 HQAKQGALNLQWT 681
L+L WT
Sbjct: 923 -----LSLPWT 928

tr Q7WQ80 Autotransporter [BB0452] [Bordetella bronchiseptica 2377
(Alcaligenes AA
bronchisepticus)] align

Score = 53.9 bits (128), Expect = 1e-05

Identities = 58/257 (22%), Positives = 103/257 (39%), Gaps = 29/257 (11%)

Query: 408 KGLWLRLVIDGHSNQWVQGTAP-VEGYRKGVLGGEVFTWQNESNQL-SIGLMGGQAEQR 465
+G W R + G+ AP +G+ G QLG +++ + + + G+ GG A R
Sbjct: 2103 RGGWARTFGRRRFERSAGGEAAPSFDGHLAQAQLGADLYARSSATRHTDAFGVFGGYATVR 2162

Query: 466 STFHN---PDNDLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFRRHINTEDG 522
H + + T ++ G Y T G Y D+ + R+R + +
Sbjct: 2163 GDVHGLARGEIQAVGTSTLRATQLGAYWT----HTGPGGWYIDTVLAGTRYRQQTSSAQ 2218

Query: 523 TERFTSK-GITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGKFSDSENAHV 581
+ G+TAS+EAGY L + R ++PQAQ+ Y + + V
Sbjct: 2219 VGAVSRGWGMTASVEAGYPWQL-----NPRWRIEPQAQVVYQQLGIANGADRSTV 2269

Query: 582 NLLGSRQLQTRVGVQAKAQFSLYKNIAIEPFAAVNALY-----HNKPFGVEMDGER 632
+ L R+G + Q++ Y + PF V+ L+ GV +
Sbjct: 2270 SYKTPDALTRLGTRLGQYA-YGKAQLRPFMGVSLLDHFTGADTVTFAGAHGVRASRQN 2328

Query: 633 RVINNKTAEIESQLGVAV 649
++ K +++QLG +V
Sbjct: 2329 TAVDLKAGVDTQLGKSV 2345

tr Q7WHU5 Autotransporter [BB3111] [Bordetella bronchiseptica 528
(Alcaligenes AA
bronchisepticus)] align

Score = 53.5 bits (127), Expect = 1e-05

Identities = 55/220 (25%), Positives = 89/220 (40%), Gaps = 42/220 (19%)

Query: 478 TGNVKGFGAGVYATWHQLQDKQGTGAYADSWMQYQFRHRINTEDGTER-----FTSKGIT 532
 TG G YAT+ + G Y D ++ R+ H D R + + G+
 Sbjct: 328 TGEGDSVHVGAATYIE----DGGFYMDGIVRVARYAHEFKAPDAKGRRVRGKYRANGVG 383

Query: 533 ASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLT-YLGVNGKFSDSENAHVNLLGSRQLQT 591
 AS+E G + FT G Y++PQ ++ + G + ++ V G L
 Sbjct: 384 ASLELG-----KRFTWAG---AWYVEPQLEMAVFHAQGGSYRAGQDLRVKDNGLTSLLG 434

Query: 592 RVGVQAKAQFSLYKNIAIEPFPAVNALYHNKPFVEMDGERRV----INNKTAI-----E 642
 RVG+ QF L ++P+A ++ L E DG+ V + +K+ + E
 Sbjct: 435 RVGLHVGRQFDLGDGRVVQPYAKLSWLQ-----EFDGKNTVRTNGVRHKSRLDGGRAE 487

Query: 643 SQLGVAVKIKSHLTLQATFNRTGKHHQAKQGALNLQWTF 682
 LGVA ++ H +L G + AK + WTF
 Sbjct: 488 LDLGVAAQLGKHGSLY-----GSYEYAKGSRQTMPWTF 520

tr Q7WC76 Autotransporter [BPP0452] [Bordetella 1769 AA
 parapertussis] align

Score = 53.5 bits (127), Expect = 1e-05

Identities = 58/257 (22%), Positives = 103/257 (39%), Gaps = 29/257 (11%)

Query: 408 KGLWLRVIDGHSNQWVQKGTAP-VEGYRKGVLGGEVFTWQNESNQL-SIGLMGGQAEQR 465
 +G W R + G+ AP +G+ G QLG +++ + + + G+ GG A R
 Sbjct: 1495 RGGWARTFGRRFERSAGGEAAPSFDGHLAQAQLGADLYARSSGTRHTDAFGVFGGYATVR 1554

Query: 466 STFHN---PDTDNLTTGNVKGFGAGVYATWHQLQDKQGTGAYADSWMQYQFRHRINTEDG 522
 H + + T ++ G Y T G Y D+ + R+R + +
 Sbjct: 1555 GDVHGLARGEIQAVGTSTLRATQLGAYWT---HTGPGGWYIDTVLAGTRYRQQTKSSAQ 1610

Query: 523 TERFTSK-GITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGKFSDSENAHV 581
 + G+TAS+EAGY L + R ++PQAQ+ Y + + V
 Sbjct: 1611 VGAVSRRGWMTASVEAGYPWQL-----NPRWRIEPQAQVVYQQLGIANGADRVSTV 1661

Query: 582 NLLGSRQLQTRVGVQAKAQFSLYKNIAIEPFPAVNALY-----HNKPFVEMDGER 632
 + L R+G + Q++ Y + PF V+ L+ GV +
 Sbjct: 1662 SYKTPDALTARLGTRLGQYA-YGKAQLRPFMGVSLLHDFTGADTVTFAGAHGVRASRQN 1720

Query: 633 RVINNKTAIESQLGVAV 649
 ++ K +++QLG +V
 Sbjct: 1721 TAVDLKAGVDTQLGKSV 1737

tr Q8VV95 Vag8 protein (Autotransporter) [vag8-2] [Bordetella pertussis] 915 AA
 align

Score = 52.0 bits (123), Expect = 4e-05

Identities = 61/252 (24%), Positives = 100/252 (39%), Gaps = 42/252 (16%)

Query: 436 GVQLGGEVFTWQNESNQLSIGLMGGQAEQRSTFHNPDNDLTTGNVKGFGAGVYATWHQL 495
 G++LG + + S+G++ G +E R D G+V G Y ++ L
 Sbjct: 678 GIELGLDRRVAGGATTAWSVGMLAGYSETRR-----DGGAYRAGHVHSAHVGAIVSY--L 730

Query: 496 QDKQTGAYADSWMQYQFRH--RINTED---GTERFTSKGITASIEAGYNALLAEHFTKK 550
D +G+Y D ++Y RFRH I T D + S G+ A + G + +
Sbjct: 731 ND--SGSYVDGVVKYNRFRHGFDIRTTDLKRVDAKHRSHGLGALLRGRRIDIDGGW--- 785

Query: 551 GNSLRVYLQPQAQLTYLGVNG-KFSDSENHAVNLLGSRQLQTRVGVQAKAQFSLYKNIAI 609
Y++PQA + + G ++ S V G+ R G +A Q L +
Sbjct: 786 -----YVEPQASVAWFHAGGSRYEASNGLRVRADGAHSWVLRAGAEAGRQMLANGNIV 839

Query: 610 EPFA-----AVNALYHNKPFGEVMDGERRVINNKTAI-ESQLGVAVKIKSHLTLQA 659
EP+A A NA+Y N G R V ++ E+++GV + L A
Sbjct: 840 EPYARLGWAQELGADNAVYTN-----GIRHVTRSRGGFAEARVGVGALLGKRHALYA 891

Query: 660 TFNRQTGKHHQA 671
+ G +A
Sbjct: 892 DYEYAKGARFEA 903

tr O66044 Vag8 [vag-8] [Bordetella pertussis] 915 AA
align

Score = 52.0 bits (123), Expect = 4e-05
Identities = 61/252 (24%), Positives = 100/252 (39%), Gaps = 42/252 (16%)

Query: 436 GVQLGGEVFTWQNESNQLSIGLMGGQAEQRSTFHNPDNDLTTGNVKGFGAGVYATWHQL 495
G++LG + + S+G++ G +E R D G+V G Y ++ L
Sbjct: 678 GIELGLDRRVAGGATTAWSVGMLAGYSETRR-----DGGAYRAGHVHSAHVGAIVSY--L 730

Query: 496 QDKQTGAYADSWMQYQFRH--RINTED---GTERFTSKGITASIEAGYNALLAEHFTKK 550
D +G+Y D ++Y RFRH I T D + S G+ A + G + +
Sbjct: 731 ND--SGSYVDGVVKYNRFRHGFDIRTTDLKRVDAKHRSHGLGALLRGRRIDIDGGW--- 785

Query: 551 GNSLRVYLQPQAQLTYLGVNG-KFSDSENHAVNLLGSRQLQTRVGVQAKAQFSLYKNIAI 609
Y++PQA + + G ++ S V G+ R G +A Q L +
Sbjct: 786 -----YVEPQASVAWFHAGGSRYEASNGLRVRADGAHSWVLRAGAEAGRQMLANGNIV 839

Query: 610 EPFA-----AVNALYHNKPFGEVMDGERRVINNKTAI-ESQLGVAVKIKSHLTLQA 659
EP+A A NA+Y N G R V ++ E+++GV + L A
Sbjct: 840 EPYARLGWAQELGADNAVYTN-----GIRHVTRSRGGFAEARVGVGALLGKRHALYA 891

Query: 660 TFNRQTGKHHQA 671
+ G +A
Sbjct: 892 DYEYAKGARFEA 903

tr Q7WQ83 Autotransporter [BB0450] [Bordetella bronchiseptica] 2152
(Alcaligenes AA
bronchisepticus)] align

Score = 51.6 bits (122), Expect = 5e-05
Identities = 47/220 (21%), Positives = 88/220 (39%), Gaps = 23/220 (10%)

Query: 407 RKGLWLRLVIDGHSNQWVQKTAP-VEGYRKGVQLGGEVFTWQNESNQLSI-GLMGGQAEQ 464
R+G W R + G P +G+ G+Q G +++ ++ + G+ GG A
Sbjct: 1876 RRGAWTRAFAFGRTFERSGSGDVPDFDGHVAGLQAGVDLYARRSDQGHADLAGVFGGYANA 1935

Query: 465 RSTFHNPDNDLTTGNVKGFGA-----GVYATWHQLQDKQTGAYADSWMQYQRFHRI-N 518
 R D G + Y + +G Y D+ + R+ + +
 Sbjct: 1936 RGHM-----DGFARGETGAYAGKPDNLNAYYIGGYWTHIGPSGWYVDAVLGTRYEQKAKS 1990

Query: 519 TEDGTERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGKFSDSEN 578
 + D G+TAS+EAGY + E + +++PQAQL Y + + +
 Sbjct: 1991 SNDLRTEAKGWGVTASVEAGYPVPIGEKW-----HIEPQAQLVYQRLTVSNGEDDV 2041

Query: 579 AHVNLLGSRQLQTRVGQAKAQFSLYKNIAIEPFPAVNAL 618
 + V+ + R+G + Q++ Y + PF V+ L
 Sbjct: 2042 SSVSYRTPDSVTARLGARLSGQYA-YNTTQLRPFMEVSL 2080

tr Q7WC79 Autotransporter [BPP0449] [Bordetella 1616 AA
 parapertussis] align

Score = 51.6 bits (122), Expect = 5e-05

Identities = 47/220 (21%), Positives = 88/220 (39%), Gaps = 23/220 (10%)

Query: 407 RKGLWLRVIDGHSNQWVQKTAP-VEGYRKGVLGGEVFTWQNESNQLSI-GLMGGQAEQ 464
 R+G W R + G P +G+ G+Q G +++ +++ + G+ GG A
 Sbjct: 1340 RRGAWTRAFGRTFERSGSGDVPFRFDGHVAGLQAGVDLYARRSDQGHADLAGVFGGYANA 1399

Query: 465 RSTFHNPDNDLTTGNVKGFGA-----GVYATWHQLQDKQTGAYADSWMQYQRFHRI-N 518
 R D G + Y + +G Y D+ + R+ + +
 Sbjct: 1400 RGHM-----DGFARGETGAYAGKPDNLNAYYIGGYWTHIGPSGWYVDAVLGTRYEQKAKS 1454

Query: 519 TEDGTERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGKFSDSEN 578
 + D G+TAS+EAGY + E + +++PQAQL Y + + +
 Sbjct: 1455 SNDLRTEAKGWGVTASVEAGYPVPIGEKW-----HIEPQAQLVYQRLTVSNGEDDV 1505

Query: 579 AHVNLLGSRQLQTRVGQAKAQFSLYKNIAIEPFPAVNAL 618
 + V+ + R+G + Q++ Y + PF V+ L
 Sbjct: 1506 SSVSYRTPDSVTARLGARLSGQYA-YNTTQLRPFMEVSL 1544

tr Q7W9Y1 Autotransporter [BPP1618] [Bordetella parapertussis] 519 AA
align

Score = 51.6 bits (122), Expect = 5e-05

Identities = 57/243 (23%), Positives = 92/243 (37%), Gaps = 47/243 (19%)

Query: 455 IGLMGGQAEQRSTFHNPDNDLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRF 514
 +G + G R F D TG+ G YAT+ + G Y D ++ R R
 Sbjct: 301 VGAVAGYTNGRIKF-----DRGGTGDDDSVHVGAATYIE----DGGFYMDGIVRVSRIR 351

Query: 515 HRINTEDGTER-----FTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGV 569
 H +D R + G+ AS+E G + FT G Y++PQ ++
 Sbjct: 352 HAFKVDDAKGRRVRGQYRGNGVGASLELG-----KRFTWPG---AWYVEPQLEVAAFRA 402

Query: 570 NG-KFSDSENAHVNNLLGSRQLQTRVGQAKAQFSLYKNIAIEPFPAVNALYHNKPFGVEM 628
 G ++ S + G+ + R+G+ QF L ++P+ ++ + E
 Sbjct: 403 QGADYTASNGLRIKDDGTNSMLGRLGLHVGRQFDLGDGRVVQPYMKLSWVQ-----EF 455

Query: 629 DGERRVINNKT-----AIESQLGVAVKIKSHLTLQATFNROTGKHHQAKQGALNLQ 679
 DG+ V N E LGVA ++ H +L G + AK +
 Sbjct: 456 DGKGTVRTNGVRHKNRLDGGRAELDLGVAAQLGKHGSLY-----GSYEYAKGSRQTMP 508

Query: 680 WTF 682
 WTF
 Sbjct: 509 WTF 511

tr Q9F290 **YapC protein (Putative autotransporter) (Putative autotransporter protein) [yapC] [Yersinia pestis]** 638
 AA
align

Score = 51.2 bits (121), Expect = 6e-05
 Identities = 39/205 (19%), Positives = 90/205 (43%), Gaps = 22/205 (10%)

Query: 409 GLWLRVIDGHSNQWVQGKTAPVEGYRKGVLGGEVFTWQNESNQLSIGLMGGQAEQRSTF 468
 G+W R + S + K + G++G + N N L + +
 Sbjct: 382 GVWARYLTDDS-RLSDNKNI AFKNTLSGMEIGADKQLGLNRGNMLIGAFTSYSSSDVKST 440

Query: 469 HNPDTDNLTGTVNKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFHRINTEDGTERFTS 528
 H+ + G+++ +G G+Y T+ Q+G Y D+ ++ RF +++NT++ +
 Sbjct: 441 HDAN-----GDIRSYGGGLYLTLYLD---QSGFYVDTVLKANRFNNKMNTQETRGEYNQ 490

Query: 529 KGITASIEAGYNALLAEHFTKKGNLSRVYLQPQAQLTYLGV-NGKFSDSENAHVNLLGSR 587
 +T S+E+GY + + + L+P +++Y + + ++ S + +
 Sbjct: 491 NALTTSVESGYQWPVYAN-----LVLEPYGKVSYSRIGSADYTLNMGVAEVAKAD 541

Query: 588 QLQTRVGVOAKAQFSLYKNIAIEPF 612
 +Q +G A +S+ + I+P+
 Sbjct: 542 SVQGEGLGTVLAASYSI-NQMTIKPY 565

tr Q88ID0 **Outer membrane autotransporter [PP3069] [Pseudomonas putida (strain KT2440)]** 825
 AA
align

Score = 50.8 bits (120), Expect = 8e-05
 Identities = 57/222 (25%), Positives = 97/222 (43%), Gaps = 36/222 (16%)

Query: 477 TTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFHRINTEDG-TERFTSKGITASI 535
 + G++K G V W + + GAY D+ +QY R R +E G T TAS+
 Sbjct: 624 SVGDLKLDGDSVGTYWTLVGPPQ--GAYLDAVLQYTRLDGRARSERGDTLNLGDHAWTASL 681

Query: 536 EAGYNALLAEHFTKKGNLSRVYLQPQAQLTYLGVNGKFSDSENAHVNLLGSRQLQTRVG 595
 E+GY L+E R ++PQAQL V + + + ++ +L R+G+
 Sbjct: 682 ESGYPITLSE-----RWRVEPQAQLIAQKVALESARDSVSRISHDAQVELTGRLGL 732

Query: 596 QAKAQFSLYKNIAIEPFPAVNALYHNKPFGEVMDGERRVIN-----NKTAIESQ 644
 + + F+ ++PFA VN L+H DG R + T+++ +
 Sbjct: 733 RLEGFTGSSGRLLQPFAQVN-LWHG-----DGGRDTLTFDDADKIKTDYRYTSVQLE 784

Query: 645 LGVAVKIKSHLTL---QATFNROTGKHHQAKQGALNLQWTF 682
 GV ++ L+L Q T N + + +A L ++W F
 Sbjct: 785 SGVVAQVNEALSLHGGVQYTANLDS-RQQEASGVNLGVRWQF 825

sp P76017 Hypothetical protein ycgV [ycgV] [Escherichia coli] 955 AA
YCGV_ECOLI align

Score = 50.1 bits (118), Expect = 1e-04

Identities = 58/218 (26%), Positives = 94/218 (42%), Gaps = 32/218 (14%)

Query: 410 LWLRVIDGHSNQWVQGKTAPVEGYRKGVLGGEVFTWQNESNQLSIGLMGGQAEQRSTFH 469
+WLR G + + GK + + G+Q GG+ ++ L +GL G ST

Sbjct: 694 IWLRSYGGSLDSFASGKLSGFDMGYSIGQFGGD--KRLSDVMPLYVGLYIG-----STHA 746

Query: 470 NPDTDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQFRHRINTEDG-----TE 524
+PD G + G+YA++ Q G Y+D ++ R ++ + D

Sbjct: 747 SPDYSG-GDGTARSDYMGMYASYM----AQNGFYSDLVIKASRQKNSFHVLDSONNGVNA 801

Query: 525 RFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNG---KFSDSENAHV 581
T+ G++ S+EAG + F Y++PQ QLTY N K S+ N H+

Sbjct: 802 NGTANGMSISLEAG-----QRFNLSPTGYGFYIEPQTQLTYSHQNEMTMKASNGLNIHL 855

Query: 582 N---LLGSRQLQTRVGVA-KAQFSLY-KNIAIEPFA 613

N LLG + + A +Q ++Y K AI F+

Sbjct: 856 NHYESLLGRASMILGYDITAGNSQLNVYVKTGAIREFS 893

tr Q881W9 Autotransporter, putative [PSPT02763] [Pseudomonas syringae] 927
(pv. tomato)] AA align

Score = 49.7 bits (117), Expect = 2e-04

Identities = 69/290 (23%), Positives = 112/290 (37%), Gaps = 39/290 (13%)

Query: 411 WLRVIDGHSNQWVQGKTAP-VEGYRKGVLGGEVFTWQNESNQ-LSIGLMGGQAEQ---- 464
W RV Q G +P ++ KG Q+G +V+ W + Q L IGL Q

Sbjct: 657 WARVFGSDFKQWQSGTVSPGLDASLKGQIGHDVYAWSLDGQQILRIGLFVAQNRLDGKV 716

Query: 465 ---RSTFHNPDTDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQFRHRINTED 521
FH TG +K G V A W T +Y D+ + R ++

Sbjct: 717 QGFAGGFHARH-----TGRIKLHGDSVGAYW--TLSSPTASYVDALVMSTRLDGYSRSDR 769

Query: 522 GTERFTSKG--ITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGKFSDSENA 579
G R ++G ++ S+EAG+ +L R +PQ Q+ + ++ +

Sbjct: 770 GL-RIDTQGHALSLSVEAGHPFVLTP-----RWVAEPQVQIIHQRIDLDDQHDGIS 819

Query: 580 HVNLLGSRQLQTRVGVAKAQFSLYKNIAIEPFAAVNALYHNK--PFGVEMDGERRVINN 637
HV R+G++ K +++L + IEP+ N L+ N V D R+

Sbjct: 820 HVGFDSPYNTGRLGIRFKGRYAL-AGMPIEPYLRAN-LWRNAGGHDTVTFDHTERIKTA 877

Query: 638 KTAIESQLGVAVKIK----SHLTLQATFNROTGKHHQAKQGA---LNLQW 680

+ LG + IK + + A +NR H + A + L W

Sbjct: 878 HRSTTGSLGAGMVIKVASDTSVYWGADYNRDLNSHDSSGANASLGVRLAW 927

tr Q7WL85 Autotransporter [vag8] [*Bordetella bronchiseptica* (*Alcaligenes bronchisepticus*)]

align

Score = 49.7 bits (117), Expect = 2e-04

Identities = 59/252 (23%), Positives = 100/252 (39%), Gaps = 42/252 (16%)

Query: 436 GVQLGGEVFTWQNESNQLSIGLMGGQAEQRSTFHNPDTDNLTGTVKGFAGVYATWHQL 495
G++LG + + S+G++ G +E R D G+V G Y ++ L
Sbjct: 678 GIELGIDRRVAGGATTAWSVGMLAGYSETRR-----DGGAYRAGHVHSAHVGAYVSY--L 730

Query: 496 QDKQTGAYADSWMQYQFRH--RINTED---GTERFTSKGITASIEAGYNALLAEHFTKK 550
D +G+Y D ++Y RFRH + T D + S G+ A + G + +
Sbjct: 731 ND--SGSYVDGVVKYNRFRHGFDVRTTDLKRVDKHRSHGLGALLRGRRIDIDGGW--- 785

Query: 551 GNSLRVYLQPQAQLTYLGVNG-KFSDSENAHVNLLGSRQLQTRVGVQAKAQFSLYKNIAI 609
Y++PQA + + +G ++ S V + R G +A Q L +
Sbjct: 786 -----YVEPQASVAWFHADGSRYEASNGLRVRADSAHSWVLRAGAEAGRQMRLANGNIV 839

Query: 610 EPFA-----AVNALYHNKPFGEVMDGERRVINNKTAI-ESQLGVAVKIKSHLTLQA 659
EP+A A NA+Y N G R V ++ E+++GV + L A
Sbjct: 840 EPYARLGWAQELGADNAVYTN-----GIRHVTRSRGGFAEARVGVGALLGKRHALYA 891

Query: 660 TFNRQTGKHHQA 671
+ G +A
Sbjct: 892 DYEYAKGARFEA 903

tr Q7W7U5 Autotransporter [vag8] [*Bordetella parapertussis*] 915 AA
align

Score = 49.7 bits (117), Expect = 2e-04

Identities = 59/252 (23%), Positives = 100/252 (39%), Gaps = 42/252 (16%)

Query: 436 GVQLGGEVFTWQNESNQLSIGLMGGQAEQRSTFHNPDTDNLTGTVKGFAGVYATWHQL 495
G++LG + + S+G++ G +E R D G+V G Y ++ L
Sbjct: 678 GIELGIDRRVAGGATTAWSVGMLAGYSETRR-----DGGAYRAGHVHSAHVGAYVSY--L 730

Query: 496 QDKQTGAYADSWMQYQFRH--RINTED---GTERFTSKGITASIEAGYNALLAEHFTKK 550
D +G+Y D ++Y RFRH + T D + S G+ A + G + +
Sbjct: 731 ND--SGSYVDGVVKYNRFRHGFDVRTTDLKRVDKHRSHGLGALLRGRRIDIDGGW--- 785

Query: 551 GNSLRVYLQPQAQLTYLGVNG-KFSDSENAHVNLLGSRQLQTRVGVQAKAQFSLYKNIAI 609
Y++PQA + + +G ++ S V + R G +A Q L +
Sbjct: 786 -----YVEPQASVAWFHADGSRYEASNGLRVRADSAHSWVLRAGAEAGRQMRLANGNIV 839

Query: 610 EPFA-----AVNALYHNKPFGEVMDGERRVINNKTAI-ESQLGVAVKIKSHLTLQA 659
EP+A A NA+Y N G R V ++ E+++GV + L A
Sbjct: 840 EPYARLGWAQELGADNAVYTN-----GIRHVTRSRGGFAEARVGVGALLGKRHALYA 891

Query: 660 TFNRQTGKHHQA 671
+ G +A
Sbjct: 892 DYEYAKGARFEA 903

tr Q7WHU6 Autotransporter [BB3110] [*Bordetella bronchiseptica* 770
(*Alcaligenes* AA
bronchisepticus)] align

Score = 49.3 bits (116), Expect = 2e-04

Identities = 57/243 (23%), Positives = 95/243 (38%), Gaps = 47/243 (19%)

Query: 455 IGLMGGQAEQRSTFHNPDNDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFR 514
+G + G R F D TG+ G YAT+ + G Y D ++ R R
Sbjct: 552 VGAVAGYTNRIKF-----DRGGTGDDDSVHVGAYATYIE----DGGFYMDGIVRVSRIR 602

Query: 515 HRINTEDGTER-----FTSKGITASIEAGYNALLAEHFTKKGNLSRVYLQPQAQLTYLGV 569
H +D R + G+ AS+E G + FT G Y++PQ ++
Sbjct: 603 HAFKVDDAKGRRVRGQYRGNGVGASLELG-----KRFTWPG---AWYVEPQLEVAAFHA 653

Query: 570 NG-KFSDSENAHVNLGSRQLQTRVGVQAKAQFSLYKNIAIEPFPAAVNALYHNKPFQVEM 628
G ++ S + G+ + R+G+ QF L ++P+ ++ + E
Sbjct: 654 QGADYTASNGLRIKDDGTNSMLGRLGLHVGRQFDLGDGRVVQPYMKLSWVQ-----EF 706

Query: 629 DGERRV----INNKTAI-----ESQLGVAVKIKSHLTQATFNROTGKHHQAKQGALNLQ 679
DG+ V I +K + E +GVA ++ H +L G + AK +
Sbjct: 707 DGKGTVRTNDRHKVRLDGGRAELAVGVASQLGKHGSL-----FGSYEYAKGSRQTMP 759

Query: 680 WTF 682
WTF
Sbjct: 760 WTF 762

tr O86135 Tracheal colonization factor precursor [tcfA2] [*Bordetella* 647 AA
pertussis] align

Score = 49.3 bits (116), Expect = 2e-04

Identities = 57/243 (23%), Positives = 95/243 (38%), Gaps = 47/243 (19%)

Query: 455 IGLMGGQAEQRSTFHNPDNDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFR 514
+G + G R F D TG+ G YAT+ + G Y D ++ R R
Sbjct: 429 VGAVAGYTNRIKF-----DRGGTGDDDSVHVGAYATYIE----DGGFYMDGIVRVSRIR 479

Query: 515 HRINTEDGTER-----FTSKGITASIEAGYNALLAEHFTKKGNLSRVYLQPQAQLTYLGV 569
H +D R + G+ AS+E G + FT G Y++PQ ++
Sbjct: 480 HAFKVDDAKGRRVRGQYRGNGVGASLELG-----KRFTWPG---AWYVEPQLEVAAFHA 530

Query: 570 NG-KFSDSENAHVNLGSRQLQTRVGVQAKAQFSLYKNIAIEPFPAAVNALYHNKPFQVEM 628
G ++ S + G+ + R+G+ QF L ++P+ ++ + E
Sbjct: 531 QGADYTASNGLRIKDDGTNSMLGRLGLHVGRQFDLGDGRVVQPYMKLSWVQ-----EF 583

Query: 629 DGERRV----INNKTAI-----ESQLGVAVKIKSHLTQATFNROTGKHHQAKQGALNLQ 679
DG+ V I +K + E +GVA ++ H +L G + AK +
Sbjct: 584 DGKGTVRTNDRHKVRLDGGRTTELAVGVASQLGKHGSL-----FGSYEYAKGSRQTMP 636

Query: 680 WTF 682
WTF
Sbjct: 637 WTF 639

Score = 49.3 bits (116), Expect = 2e-04

Identities = 57/243 (23%), Positives = 95/243 (38%), Gaps = 47/243 (19%)

Query: 455 IGLMGGQAEQRSTFHNPDNDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFR 514
+G + G R F D TG+ G YAT+ + G Y D ++ R R

Sbjct: 429 VGAVAGYTNGRIKF-----DRGGTGDDDSVHVGAYATYIE----DGGFYMDGIVRVSRIR 479

Query: 515 HRINTEDGTER-----FTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGV 569
H +D R + G+ AS+E G + FT G Y++PQ ++

Sbjct: 480 HAFKVDDAKGRRVRGQYRGNGVGASLELG-----KRFTWPG---AWYVEPQLEVAAFHA 530

Query: 570 NG-KFSDSENAHVNLGSRQLQTRVGVOAKAQFSLYKNIAIEPFPAVNALYHNKPFGVEM 628
G ++ S + G+ + R+G+ QF L ++P+ ++ + E

Sbjct: 531 QGADYTASNGLRIKDDGTNSMLGRLGLHVGRQFDLGDGRVVQPYMKLSWVQ-----EF 583

Query: 629 DGERRV----INNKTAI-----ESQLGVAVKIKSHLTLQATFNRTGKHHQAKQGALNLQ 679
DG+ V I +K + E +GVA ++ H +L G + AK +

Sbjct: 584 DGKGTVRTNDIRHKVRLDGGRTTELAVGVASQLGKHGSL-----FGSYEYAKGSRQTMP 636

Query: 680 WTF 682

WTF

Sbjct: 637 WTF 639

tr Q9F4B3 BapB protein [bapB] [Bordetella pertussis] 482 AA
align

Score = 49.3 bits (116), Expect = 2e-04

Identities = 57/243 (23%), Positives = 95/243 (38%), Gaps = 47/243 (19%)

Query: 455 IGLMGGQAEQRSTFHNPDNDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFR 514
+G + G R F D TG+ G YAT+ + G Y D ++ R R

Sbjct: 264 VGAVAGYTNGRIKF-----DRGGTGDDDSVHVGAYATYIE----DGGFYMDGIVRVSRIR 314

Query: 515 HRINTEDGTER-----FTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGV 569
H +D R + G+ AS+E G + FT G Y++PQ ++

Sbjct: 315 HAFKVDDAKGRRVRGQYRGNGVGASLELG-----KRFTWPG---AWYVEPQLEVAAFHA 365

Query: 570 NG-KFSDSENAHVNLGSRQLQTRVGVOAKAQFSLYKNIAIEPFPAVNALYHNKPFGVEM 628
G ++ S + G+ + R+G+ QF L ++P+ ++ + E

Sbjct: 366 QGADYTASNGLRIKDDGTNSMLGRLGLHVGRQFDLGDGRVVQPYMKLSWVQ-----EF 418

Query: 629 DGERRV----INNKTAI-----ESQLGVAVKIKSHLTLQATFNRTGKHHQAKQGALNLQ 679
DG+ V I +K + E +GVA ++ H +L G + AK +

Sbjct: 419 DGKGTVRTNDIRHKVRLDGGRTTELAVGVASQLGKHGSL-----FGSYEYAKGSRQTMP 471

Query: 680 WTF 682

WTF

Sbjct: 472 WTF 474

tr Q8GB90 Tracheal colonization factor [tcfA] [Bordetella pertussis] 647 AA

align

Score = 49.3 bits (116), Expect = 2e-04

Identities = 57/243 (23%), Positives = 95/243 (38%), Gaps = 47/243 (19%)

Query: 455 IGLMGGQAEQRSTFHNPDNDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRF 514
+G + G R F D TG+ G YAT+ + G Y D ++ R R
Sbjct: 429 VGAVAGYTNGRIKF-----DRGGTGDDDSVHVGAAYATYIE---DGGFYMDGIVRVSRIR 479

Query: 515 HRINTEDGTER-----FTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGV 569
H +D R + G+ AS+E G + FT G Y++PQ ++
Sbjct: 480 HAFKVDDAKGRRVRGQYRGNGVGASLELG-----KRFTWPG---AWYVEPQLEVAAFHA 530

Query: 570 NG-KFSDSENAHVNLGSRQLQTRVGVQAKAQFSLYKNIAIEPFAAVNALYHNKPFVEM 628
G ++ S + G+ + R+G+ QF L ++P+ ++ + E
Sbjct: 531 QGADYTASNGLRIKDDGTNSMLGRLGLHVGRQFDLGDGRVVQPYMKLSWVQ-----EF 583

Query: 629 DGERRV----INNKTAI-----ESQLGVAVKIKSHLTLQATFNRTGKHHQAKQGALNLQ 679
DG+ V I +K + E +GVA ++ H +L G + AK +
Sbjct: 584 DGKGTVRTNDRHKVRLDGGRTTELAVGVASQLGKHGSL-----FGSYEYAKGSRQTMP 636

Query: 680 WTF 682
WTF
Sbjct: 637 WTF 639

tr Q45343 Tracheal colonization factor [tcfA] [Bordetella pertussis] 672 AAalign

Score = 49.3 bits (116), Expect = 2e-04

Identities = 57/243 (23%), Positives = 95/243 (38%), Gaps = 47/243 (19%)

Query: 455 IGLMGGQAEQRSTFHNPDNDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRF 514
+G + G R F D TG+ G YAT+ + G Y D ++ R R
Sbjct: 454 VGAVAGYTNGRIKF-----DRGGTGDDDSVHVGAAYATYIE---DGGFYMDGIVRVSRIR 504

Query: 515 HRINTEDGTER-----FTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGV 569
H +D R + G+ AS+E G + FT G Y++PQ ++
Sbjct: 505 HAFKVDDAKGRRVRGQYRGNGVGASLELG-----KRFTWPG---AWYVEPQLEVAAFHA 555

Query: 570 NG-KFSDSENAHVNLGSRQLQTRVGVQAKAQFSLYKNIAIEPFAAVNALYHNKPFVEM 628
G ++ S + G+ + R+G+ QF L ++P+ ++ + E
Sbjct: 556 QGADYTASNGLRIKDDGTNSMLGRLGLHVGRQFDLGDGRVVQPYMKLSWVQ-----EF 608

Query: 629 DGERRV----INNKTAI-----ESQLGVAVKIKSHLTLQATFNRTGKHHQAKQGALNLQ 679
DG+ V I +K + E +GVA ++ H +L G + AK +
Sbjct: 609 DGKGTVRTNDRHKVRLDGGRTTELAVGVASQLGKHGSL-----FGSYEYAKGSRQTMP 661

Query: 680 WTF 682
WTF
Sbjct: 662 WTF 664

tr Q6U948 **Tracheal colonization factor protein [tcfA] [Bordetella pertussis]**

642
AA
align

Score = 49.3 bits (116), Expect = 2e-04

Identities = 57/243 (23%), Positives = 95/243 (38%), Gaps = 47/243 (19%)

Query: 455 IGLMGGQAEQRSTFHNPDNDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRF 514
+G + G R F D TG+ G YAT+ + G Y D ++ R R
Sbjct: 424 VGAVAGYTNGRIKF-----DRGGTGDDDSVHVGAYATYIE----DGGFYMDGIVRVSRIR 474

Query: 515 HRINTEDGTER-----FTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGV 569
H +D R + G+ AS+E G + FT G Y++PQ ++
Sbjct: 475 HAFKVDDAKGRRVRGQYRGNVGASLELG-----KRFTWPG---AWYVEPQLEVAAFHA 525

Query: 570 NG-KFSDSENAHVNLLGSRQLQTRVGVQAKAQFSLYKNIAIEPFAAVNALYHNKPFQGVEM 628
G ++ S + G+ + R+G+ QF L ++P+ ++ + E
Sbjct: 526 QGADYTASNGLRIKDDGTNSMLGRLGLHVGRQFDLGDGRVVQPYMKLSWVQ-----EF 578

Query: 629 DGERRV----INNKTAI-----ESQLGVAVKIKSHLTLQATFNRTGKHHQAKQGALNLQ 679
DG+ V I +K + E +GVA ++ H +L G + AK +
Sbjct: 579 DGKGTVRTNDIRHKVRLDGGRTAVGASQLGKHGSL-----FGSYEYAKGSRQTMP 631

Query: 680 WTF 682

WTF

Sbjct: 632 WTF 634

tr Q7WKS2 **Putative autotransporter [bapC] [Bordetella bronchiseptica 993 AA (Alcaligenes bronchisepticus)]**

align

Score = 48.1 bits (113), Expect = 5e-04

Identities = 96/486 (19%), Positives = 174/486 (35%), Gaps = 80/486 (16%)

Query: 234 GQTISLDEFRLNSLWEPWDSNVGKLKTTNADIRFNTKSES-----LLVKEDYAGGAR 287
G T +++ R+E+ W S V L + + T +ES L VK +G
Sbjct: 543 GITNTVNALRIEDGTWTVTGSSAVNSLHLQAGKVAYATPAESDQKFKHLRVKT-LSGSG 601

Query: 288 FRFAYDPKEAKNTALIFEKNVTGTSDIIFENP-----IDDLKSLD-----GHQIIKV 334
F + L+ +G ++ ++ L ++ G +
Sbjct: 602 FEMNASADLSDGDLVVSDEASGQHKVLRGAGTEPTGVESLTLVELPKGSQAGFTLANR 661

Query: 335 NGTADKHAFLRSLGKHQKGIYTLSQLQRPEGFLPKCKN-----AMIWRFMHNRLKPPIP 387
G D AFR G++ L + N + IW N L +
Sbjct: 662 GGVVDAGAFRYRLTPDNGVWGLERTSQLSAVANAALNTGGVGAASSIWAEGNALSKRLG 721

Query: 388 YCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGKTAPVEGYRKGVQ--LGGEVFT 445
LRL+ + RT +K + +D + + K V G+ G + G+
Sbjct: 722 --ELRLDPGAGGFWGRTFAQK----QQLDNKAGRRFDQK--VYGFEFGADHAIAGQQGR 772

Query: 446 WQNESNQLSIGLMGGQAEQRSTFHNPDNDNLTTGNVKGFGAGVYATWHQLQDKQTGAYAD 505
W +G + G R +F + + + ++ + A V G Y D
Sbjct: 773 WH-----VGGLLGYTRARRSFVDDGAGHTDSAHIGAYAAYV-----ADNGFYFD 816

Query: 506 SWMQYQRFH--RINTEDGTE---RFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQP 560
S ++ RF + + D ++ + G+ A++EAG L + + +++P

Sbjct: 817 STLRSRFFENDFTVTATDAVSIRGKYRANGVGATLEAGKRFTLHDGW-----FVEP 867

Query: 561 QAQLTYLGVNG-KFSDSENAHVNLLGSRQLQTRVGVQAKAQFSLYKNIAIEPFAAVNALY 619
Q++++ +G + + N V G R+G+ A + L K+ I+P+A ++ L

Sbjct: 868 QSEVSLPHASGGTYRAANNLSVKDEGGTSAVLRLGLAAGRRIELGKDRVIQPYATLSWLQ 927

Query: 620 HNKPFGEVMDGERRVINNKTAIESQLGVAVKIKSHLTLOATFNRRQ---TGKHHQAKQGAL 676
E G V N + + LG + + L L A R + AK L

Sbjct: 928 -----EFKGVTTVRTNGYGLRDTDLG-GGRAELALGLAAALGRGHKFYTSYEYAKGNKL 979

Query: 677 NLQWTF 682
L WTF

Sbjct: 980 TLPWTF 985

tr Q83RP3 Putative adhesion and penetration protein [SF1205] [Shigella flexneri] 955 AA
align

Score = 47.4 bits (111), Expect = 0.001

Identities = 57/218 (26%), Positives = 93/218 (42%), Gaps = 32/218 (14%)

Query: 410 LWLRVIDGHSNQWVQGKTAPVEGYRKGVQLGGEVFTWQNESNQLSIGLMGGQAEQRSTFH 469
+WLR G + + GK + + G+Q GG+ ++ L +GL ST

Sbjct: 694 IWLRSYGGSLSDFASGKLSGFDMGYSIGIQFGGD--KRLSDVMPLYVGLY-----IDSTHA 746

Query: 470 NPDTDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFRRHINTEDG-----TE 524
+PD G + G+YA++ Q G Y+D ++ R ++ + D

Sbjct: 747 SPDYSG-GDGTARSDYMGMYASYM----AQNGFYSDLVIKASRQKNSFHVLDSONNGVNA 801

Query: 525 RFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNG---KFSDSENAHV 581
T+ G++ S+EAG + F Y++PQ QLTY N K S+ N H+

Sbjct: 802 NGTANGMSISLEAG-----QRFNLSPTGYGYIEPQTQLTYSHQNEAMKASNGLNIHL 855

Query: 582 N----LLGSRQLQTRVGVQA-KAQFSLY-KNIAIEPFA 613
N LLG + + A +Q ++Y K AI F+

Sbjct: 856 NHYESLLGRASMILGYDITAGNSQLNVYVKTGAIREFS 893

tr Q7WK68 Autotransporter [BB2270] [Bordetella bronchiseptica] 559
(Alcaligenes bronchisepticus) AA
align

Score = 47.4 bits (111), Expect = 0.001

Identities = 50/219 (22%), Positives = 85/219 (37%), Gaps = 42/219 (19%)

Query: 479 GNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFRRHINTEDG-----TERFTSKGITA 533
GN G YAT+ G Y D ++ R+ H + T ++ + GI

Sbjct: 360 GNSDSLHVGAAYATYIG---DGGFYLDGIVRVNRYEHDFKADGQRGARVTGKYRANGIGL 415

Query: 534 SIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGK-FSDSENAHVNLLGSRQLQTR 592
S+E G FT G+ +++PQ ++ G ++ S V++ ++ L R

Sbjct: 416 SLETG-----RRFTWAGDW---FVEPQVEVALFRSGGADYTASNGVRVDVASTKSLGR 466

Query: 593 VGVQAKAQFSLYKNIAIEPFAAVNALYHNKPFGEVMDGERRVINNKTA-----IES 643

G+Q + L ++P+A ++ L E DG +V N E
Sbjct: 467 AGLQVGRKLDLGNKGLVQPYAKLSWLQ-----EFDGVGKVRTNDIGHDVKLGRGRAEL 519
Query: 644 QLGVAVKIKSHLTLQATFNRQTGKHHQAKQGALNLQWTF 682
LGVA + H +L A++ G L + W+F
Sbjct: 520 DLGVAAALGRHSSLFASYEYSKGSR-----LTIPWSF 551

tr Q7VVD6 Autotransporter (Pseudogene) [bapC] [Bordetella pertussis] 993 AA

align

Score = 47.4 bits (111), Expect = 0.001

Identities = 93/486 (19%), Positives = 173/486 (35%), Gaps = 80/486 (16%)

Query: 234 GQTISLDEFRLNSLWEPRWDSNVGKLKTTNADIRFNTKSES-----LLVKEDYAGGAR 287
G T +++ R+E+ W S V L + + T +ES L VK +G
Sbjct: 543 GMTKTVNALRIEDGTWTVTGSSTVNSLHLQAGKVAYATPAESDGEFKHLRVK-TLSGSGSL 601
Query: 288 FRFAYDPKEAKNTALIFEKNVTGTSDIIFENPIDDLKSLDGHQIIKV----- 334
F + L+ +G ++ + ++ +++++
Sbjct: 602 FEMNASADLSDGDLVVSDEASGQHKVLVRGAGTEPTGVESLTLVELPEGSQTKFTLANR 661
Query: 335 NGTADKHAFLRLSGKHQKGIYTLSSLQQRPEGFLPKCKN-----AMIWRFMHNRLKPPIP 387
G D AFR G++ L + N + IW N L +
Sbjct: 662 GGVVDAGAFRYRLTPDNGVWGLERTSQLSAVANAALNTGGVGAASSIWYAEGNALSKRLG 721
Query: 388 YCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQKTAPEGYRKGVQ--LGGEVFT 445
LRL+ + RT +K + +D + + K V G+ G + G+
Sbjct: 722 --ELRLDPGAGGFWRGRTFAQK---QQLDNKAGRRFDQK---VYGFELGADHAIAQQQGR 772
Query: 446 WQNESNQLSIGLMGGQAEQRSTFHNPDTDNLTTGNVKGFGAGVYATWHQLQDKQTGAYAD 505
W +G + G R +F + + + ++ + A V G Y D
Sbjct: 773 WH-----VGGLLGYTRARRSFIDDGAGHTDSAHIGAYAAAYV-----ADNGFYFD 816
Query: 506 SWMQYQRFRRH--RINTEDGTE--RFTSKGITASIEAGYNALLAEHFTKKGNLSLRVYLQP 560
S ++ RF + + D ++ + G+ A++EAG L + + +++P
Sbjct: 817 STLRSARFENDFTVTATDAVSVRGKYRANGVGATLEAGKRFTLHDGW-----FVEP 867
Query: 561 QAQLT-YLGVNGKFSDSENAHVNLLGSRQLQTRVGVAQAKAQFSLYKNIAIEPFPAVNALY 619
Q++++ + G + + N V G R+G+ A + L K+ I+P+A ++ L
Sbjct: 868 QSEVSLFHASGGTYRAANNLSVKDEGGTSAVRLRLGLAAGRRIDLKDRVIQPYATLSWLQ 927
Query: 620 HNKPFVEMDGERRVINNKTAIESQLGVAVKIKSHLTLQATFNR---QTGKHHQAKQGAL 676
E G V N + + L + + L L A R + AK L
Sbjct: 928 -----EFKGVTTVRTNGYGLRTDLS-GGRAELALGLAAALGRGHQLYTSYEYAKGNKL 979
Query: 677 NLQWTF 682
L WTF
Sbjct: 980 TLPWTF 985

tr Q7C1Y6 Putative adhesion and penetration protein [S1289] [Shigella flexneri] 955 AA

align

tr O86044 Putative autotransporter (BapC protein) [bapC] [Bordetella 759 AA
pertussis]

align

Score = 47.4 bits (111), Expect = 0.001

Identities = 93/486 (19%), Positives = 173/486 (35%), Gaps = 80/486 (16%)

Query: 234 GQTISLDEFRLNSLWEPRWDSNVGKLKTTNADIRFNTKSES-----LLVKEDYAGGAR 287
G T +++ R+E+ W S V L + + T +ES L VK +G

Sbjct: 309 GMTKTVNALRIEDGTWTVTGSSTVNSLHLQAGKVAYATPAESDGEFKHLRVK-TLSGSGSL 367

Query: 288 FRFAYDPKEAKNTALIFEKNVTGTSDIIFENPIDDLKSLDGHQIIKV----- 334
F + L+ +G ++ + ++ +++++

Sbjct: 368 FEMNASADLSDGDLVVSDEASGQHKVLVRGAGTEPTGVESLTLVELPEGSQTKFTLANR 427

Query: 335 NGTADKHAFLRLSGKHQKGIYTLSQLQRPEGFLPKCKN-----AMIWRFMHNRLKPPIP 387
G D AFR G++ L + N + IW N L +

Sbjct: 428 GGVVDAGAFRYRLTPDNGVWGLERTSQLSAVANAALNTGGVGAASSIWYAEGNALSKRLG 487

Query: 388 YCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQKGTAPVEGYRKGVQ--LGGEVFT 445
LRL+ + RT +K + +D + + K V G+ G + G+

Sbjct: 488 --ELRLDPGAGGFWRGRTFAQK---QQLDNKAGRRFDQK---VYGFELGADHAIAGQQGR 538

Query: 446 WQNESNQLSIGLMGGQAEQRSTFHNPDTDNLTTGNVKGFGAGVYATWHQLQDKQTGAYAD 505
W +G + G R +F + + + ++ + A V G Y D

Sbjct: 539 WH-----VGGLLGYTRARRSFIDDGAGHTDSAIGAYAAV-----ADNGFYFD 582

Query: 506 SWMQYQRFRRH--RINTEDGTE---RFTSKGITASIEAGYNALLAEHFTKKGNLSRVYLQP 560
S ++ RF + + D ++ + G+ A++EAG L + + +++P

Sbjct: 583 STLRSRFFENDFTVTATDAVSVRGKYRANGVGATLEAGKRFTLHDGW-----FVEP 633

Query: 561 QAQLT-YLGVNGKFSSENHAVNLLGSRQLQTRVGVQAKAQFSLYKNIAIEPFAAVNALY 619
Q++++ + G + + N V G R+G+ A + L K+ I+P+A ++ L

Sbjct: 634 QSEVSLFHASGGTYRAANNLSVKDEGGTSAVLRRLGLAAGRRIIDLKGRVIQPYATLSWLQ 693

Query: 620 HNKPFGEVMDGERRVINNKTAIESQLGVAVKIKSHLTQATFNR---QTGKHQAKQGAL 676
E G V N + + L + + L L A R + AK L

Sbjct: 694 -----EFKGVTTVRTNGYGLRTDLS-GGRAELALGLAALGRGHQLYTSYEYAKGNKL 745

Query: 677 NLQWTF 682

L WTF

Sbjct: 746 TLPWTF 751

tr Q7W8V8 Autotransporter [BPP2022] [Bordetella parapertussis] 538 AA

align

Score = 47.0 bits (110), Expect = 0.001

Identities = 52/220 (23%), Positives = 89/220 (39%), Gaps = 44/220 (20%)

Query: 479 GNVKFGAGVYATWHQLQDKQTGAYADSWMQYQRFRRHINTEDG-----TERFTSKGITA 533
GN G YAT+ G Y D ++ R+ H + T ++ + GI

Sbjct: 339 GNSDSLHVGAYATYIG----DGGFYLDGIVRVNRYEHDFKADQQRGARVTGKYRANGIGL 394

Query: 534 SIEAGYNALLAEHFTKKGNLSRVYLQQAQLTYLGVNGK--FSDSENHAVNLLGSRQLQT 591

```

      S+E G      FT  G+    +++PQ ++  L  +GK  ++ S    V++  ++ L
Sbjct: 395 SLETG-----RRFTWAGDW---FVEPQVEVA-LFRSGKADYTASNGVRVDVASTKSLLG 444

Query: 592 RVGVQAKAQFSLYKNIAIEPFAAVNALYHNKPFQVEMDGERRVINNKTA-----IE 642
      R+G+Q    +  L      ++P+A ++ L      E DG  +V  N      E
Sbjct: 445 RLGLQVGRKLDLGNGLVQPYAKLSWLQ-----EFDGVGKVRTNDIGHDVKLRRGGRAE 497

Query: 643 SQLGVAVKIKSHLTLQATFNRTGKHHQAKQGALNLQWTF 682
      LGVA  +  H +L A++    G      L  + W+F
Sbjct: 498 LDLGVAAALGKHSSLFASYEYSKGSR-----LTIPWSF 530

```

tr Q7W1F5 **Autotransporter [BPP0735] [Bordetella parapertussis]** 984 AA
align

Score = 46.2 bits (108), Expect = 0.002

Identities = 72/280 (25%), Positives = 110/280 (38%), Gaps = 39/280 (13%)

```

Query: 407 RKGLWLRVIDGHSNQWVQGTAP-VEGYRKGVLGGEVFTWQNE---SNQLSIGLMGGQA 462
      RK  W RVI    Q  G  P  +G    Q G ++    +  S++L +    GQA
Sbjct: 711 RKAGWGRVIGQRLTQRWDGDVEPRFKGNIWIAQAGADMLERDRDDGLSDRLGLFSAYGQA 770

Query: 463 EQRSTFHNPDNDLTTGN--VKGFAGVYATWHQLQDKQTGAYADSWMQYQFRHRINTE 520
      + R      G  V+ +G G+Y  W +L  K T  Y D+ +    + R  ++
Sbjct: 771 DGRVDGFVQGEHGKQAGKLRVEAYGLGLY--WTRL--KHTNWDNVLGMNYYTGRSRSD 826

Query: 521 DGTER-FTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGKFSDSENA 579
      G      G TAS EAGY+  H      + LQPQAQL Y      +++ +N
Sbjct: 827 RGVAASLEGWGFTASSEAGYS-FFPRH-----DIMLQPQAQLVY-----QYTSLDNT 872

Query: 580 H-----VNLLGSRQLQTRVGVQAKAQFSLYKNIAIEPFAAVNALYHNKPFQVEMDGERRV 634
      H      +  G  L  R+G+  +      +  I P+A +N L+H      GE
Sbjct: 873 HDAYSTIRYHGGGALTGRIGLLLQGNADQPER--IRPYARIN-LWHR-----FSHGESVS 924

Query: 635 INNKTAIESQLGVAVKIKSHLTLQATFNRTGKHHQAKQG 674
      +I ++ G +  +  + L A NRQT  +  A  G
Sbjct: 925 FGPSDSIRTEYG-STSMDLRIGLAAPLNRQTELYASAGY 963

```

tr Q7WP64 **Autotransporter [BB0821] [Bordetella bronchiseptica** 987
(Alcaligenes AA
bronchisepticus)] align

Score = 45.4 bits (106), Expect = 0.003

Identities = 71/275 (25%), Positives = 106/275 (37%), Gaps = 29/275 (10%)

```

Query: 407 RKGLWLRVIDGHSNQWVQGTAPVEGYRKGVLGGEVFTWQNESNQLS--IGLMG--GQA 462
      RK  W RVI    Q  G  P      +  G      ++  + LS  +GL G  GQA
Sbjct: 714 RKAGWGRVIGQRLTQRWDGDVEPRFKGNIWIAQAGADMLERDRDDGLSDRLGLFGAYGQA 773

Query: 463 EQRSTFHNPDNDLTTGN--VKGFAGVYATWHQLQDKQTGAYADSWMQYQFRHRINTE 520
      + R      G  V+ +G G+Y  W +L  K T  Y D+ +    + R  ++
Sbjct: 774 DGRVDGFVQGEHGKQAGKLRVEAYGLGLY--WTRL--KHTNWDNVLGMNYYTGRSRSD 829

Query: 521 DGTER-FTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGKFSDSENA 579

```

```

      G      G TAS EAGY+      H      + LQPQAQL Y      +      +
Sbjct: 830 RGVAASLEGWGF TASSEAGYS-FFPRH-----DIMLQPQAQLVYQYTSLDDTHDAYS 880

Query: 580 HVNLLGSRQLQTRVGVQAKAQFSLYKNIAIEPFAAVNALYHNKPFGVEMDGERRVINNK 639
      +      G      L R+G+      +      +      I P+A +N L+H      GE
Sbjct: 881 TIRYHGGGALTGRIGLLLQGNADQPER--IRPYARIN-LWHR-----FSHGESVSFGPSD 932

Query: 640 AIESQLGVAVKIKSHLTLQATFNRTGKHHQAKQG 674
      +I ++ G +      +      + L A NRQT + A G
Sbjct: 933 SIRTEYG-STSMDLRIGLAAPLNRTQELYASAGYG 966

```

tr Q6G430 **Hypothetical protein [BH05490] [Bartonella henselae** 1291
(Rochalimaea AA
henselae)] align

Score = 44.3 bits (103), Expect = 0.008

Identities = 55/225 (24%), Positives = 91/225 (40%), Gaps = 42/225 (18%)

```

Query: 479 GNVKG---FGAGVYATWHQLQDKQTGAYADSWMQYQFRHRINTEDGTE-----RFTSKG 530
      G V G      + G YAT+      ++G Y DS ++Y ++++ +      + G
Sbjct: 1088 GGVSGINTYSIGAYATYFD----KSGWYLDLSILKYNQYQNNLKAVSTNGIAIEGNYNQWG 1143

Query: 531 ITASIEAGYNALLAEHFTKKGNLSRVYLQPQAQLTYLGVNG---KFSDSENAHVNLLGSR 587
      + S EAGY      TK      ++QP AQL++L V G      K S+      +N S
Sbjct: 1144 LGTSFEAGYRF----ETTKSS-----WMQPYAQLSWLQVEGKEIKLSNEMTGDINAFTS- 1193

Query: 588 QLQTRVGVQAKAQFSLYKNIAIEPFAAVNALYHNKPFGVEMDGERRVINNKTAIESQL-- 645
      L++ VG+ A +F L ++      +      L N+      D R IN      + L
Sbjct: 1194 -LRSEVGLSAGYEFCLGGDVTS MAYITAAWLRENR-----DSNRTTINKLHQFVTDLSG 1246

Query: 646 -----GVAVKIKSHLTLQATFNRTG-KHHQAKQGALNLQWTF 682
      G++ +      L A      G K Q+ QG L ++++F
Sbjct: 1247 NFGKLGIGLSSLVSKKFKLYAEAQYVKGDVKVQSFQGILGVRYSF 1291

```

tr Q7W7D4 **Putative autotransporter [bapC] [Bordetella parapertussis]** 986 AA
align

Score = 44.3 bits (103), Expect = 0.008

Identities = 93/485 (19%), Positives = 172/485 (35%), Gaps = 80/485 (16%)

```

Query: 234 GQTISLDEFRLNSLWEPRWDSNVGKLTNTNADIRFNTKSES-----LLVKEDYAGGAR 287
      G T +++ R+E+ W      S V L      + + T +E+      L VK      +G
Sbjct: 536 GITKTVNALRIEDGTWTVTGSSAVNSLHLQAGKVAYATPAETNGKFKHLRVKT-LSGSG 594

Query: 288 FRFAYDPKEAKNTALIFEKNVTGTSDIIFENP-----IDDLKSLD-----GHQIIKV 334
      F      +      L+      +G      ++      ++ L ++      G +
Sbjct: 595 FEMNASADLSDGDLVVSDEASGQHKVLVRGAGTEPTGVESLTLVELPKGSQAGFTLANR 654

Query: 335 NGTADKHAFRLSGKHQKGIYTLSQLQQRPEGFLPKCKN-----AMIWRFMHNRLKPPIP 387
      G D AFR      G++ L      +      N      + IW      N L      +
Sbjct: 655 GGVVDAGAFRYRLTPDNGVWGLERTSQLSAVANAALNTGGVGAASSIWIYAKGNALSKRLG 714

```

Query: 388 YCALRLNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGKTAPVEGYRKG VQ--LGGEVFT 445
LRL+ + RT +K + +D + + K V G+ G + G+
Sbjct: 715 --ELRLDPGAGGFWGRTFAQK----QQLDNKAGRRFDQK--VYGFELGADHAIAGQQGR 765

Query: 446 WQNESNQLSIGLMGGQAEQRSTFHNPDNDNLTTGNVKGFGAGVYATWHQLQDKQTGAYAD 505
W +G + G R +F + + + ++ + A V G Y D
Sbjct: 766 WH-----VGGLLGYTRARRSFVDDGAGHTDSAIGAYAAYV-----ADNGFYFD 809

Query: 506 SWMQYQRFRR--RINTEDGTE---RFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQP 560
S ++ RF + + D ++ + G+ A++EAG L + + +++P
Sbjct: 810 STLRSRPFENDFTVTATDAVSIRGKYRANGVGATLEAGKRFTLHDGW-----FVEP 860

Query: 561 QAQLTYLGVNG-KFSDSENAHVNLGSRQLQTRVGVOAKAQFSLYKNIAIEPFAAVNALY 619
Q++++ +G + + N V G R+G+ + L K+ I+P+A ++ L
Sbjct: 861 QSEVSLFHASGGTYRAANNLSVKDEGGTSAVLRGLGLATGRRIDLGKDRVIQPYATLSWLQ 920

Query: 620 HNKPFGEVMDGERRVINNKTAIESQLGVAVKIKSHLTLQATFNRQTG---KHHQAKQ GAL 676
E G V N + + LG + + L L A R + AK L
Sbjct: 921 -----EFKGVTTVRTNGYGLRTDLG--GGRAELALGLAAALGRHSLYTSY EYAKGSKL 972

Query: 677 NLQWT 681
L WT
Sbjct: 973 TLPWT 977

tr Q9JMS5 YcbB protein [ycbB] [Escherichia coli] 1758 AA
align

Score = 43.5 bits (101), Expect = 0.013

Identities = 29/112 (25%), Positives = 54/112 (47%), Gaps = 12/112 (10%)

Query: 437 VQLGGEVFTWQNESNQLSIGLMG--QAEQRSTFHNPDNDNLTTGNVKGFGAGVYATWHQ 494
+ +GG+++ W + N IG+MG G A ++T N + TGNV G+ G+Y + Q
Sbjct: 1499 MMIGGDIYNWTDGYNYSHIGIMGGMGAANKTTSTN--NKRATGNVDGYTLGLYHV FQQ 1555

Query: 495 -----LQDKQTGAYADSWMQYQRFRRHRI-NTEDGTERFTSKGITASIEAGY 539
+ ++ G + S +QY + + + +T + + G + E GY
Sbjct: 1556 NISDGLNESERQGLWYSSIQYMDYDNSVSSTNNFKANYGVNGFRLTGEVGY 1607

tr Q9XD84 TibA [tiba] [Escherichia coli] 989 AA
align

Score = 43.1 bits (100), Expect = 0.017

Identities = 50/210 (23%), Positives = 86/210 (40%), Gaps = 26/210 (12%)

Query: 393 LNNKNSDIFDRTLPRKGLWLRVIDGHSNQWVQGKTAPVEGYRKG VQLGGEVFTWQNESNQ 452
L ++ D+ T G+W R G N+ G ++ + G + G ++ ++S+
Sbjct: 711 LRFRHGDVQMONTAPGGVWGRYT-GSDNRISGGASSGYTLTQNGFETGADMVFDLS DSS- 768

Query: 453 LSIGLMGGQAEQRSTFHNPDNDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQR 512
L++G ++ S H + NV G G+YATW G Y D ++Y R
Sbjct: 769 LAVGTFFSYSDN-SIKHARGGKS---NVDSSGGGLYATWFD---NDGYYVDGV LKYNR 819

Query: 513 FRHRINT--EDGTE---RFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYL 567

```

      F + + T   DGT      ++ G   S+EAG   L E+           + QP + T
Sbjct: 820 FNNELRTWMSDGTAVKGDYSQNGFGGSLEAGRTFSLNEN-----AWAQPYVRTTAF 870

Query: 568 GVNGKFSDSSENAHVNLLG-SRQLQTRVG VQ 596
      + K      N      +G ++ LQ   G++
Sbjct: 871 RADKKEIRLNNGMKASIGATKSLQAEAGLK 900

```

```

tr Q6G428 Hypothetical protein [BH05510] [Bartonella henselae] 874
      (Rochalimaea
      henselae)] AA
      align

```

Score = 42.4 bits (98), Expect = 0.029

Identities = 54/225 (24%), Positives = 90/225 (40%), Gaps = 42/225 (18%)

```

Query: 479 GNVKG---FGAGVYATWHQLQDKQTGAYADSWMQYQRFRRHINTEDGTE-----RFTSKG 530
      G V G   +   G YAT+      ++G Y DS ++Y ++++ +           +   G
Sbjct: 671 GGVSGINTYSIGAYATYFD---KSGWYLDLSILKYNQYQNNLKAVSTNGIAIEGNYNQWG 726

Query: 531 ITASIEAGYNALLAEHFTKKGNSLRVYLQPPAQLTYLGVNG---KFSDSENAHVNLLGSR 587
      +   S EAGY           TK      ++QP AQL++L V G   K S+      +N   S
Sbjct: 727 LGTSFEAGYRF----ETTKSS-----WMQPYAQLSWLQVEGKEIKLSNEMTGDINAFTS- 776

Query: 588 QLQTRVG VQAKAQFSLYKNIAIEPFAAVNALYHNKPFVEMDGERRVINNKTAIESQL-- 645
      L++ VG+ A   +F L   ++      +   L N+      D   IN      + L
Sbjct: 777 -LRSEVGLSAGYEFCLGGDVTSMA YITAAWLREN-----DSNHTTINKLHQFVTDLSG 829

Query: 646 -----GVAVKIKSHLTLQATFNRTG-KHHQAKQGALNLQWTF 682
      G++   +   L A           G K   Q+ QG L   ++++F
Sbjct: 830 NFGKLGIGLSSLVSKKFKLYAEAQYVKGDVKVQSFQGILGVRYSF 874

```

```

tr Q83JR2 Exported serine protease SigA [sigA] [Shigella] 1285 AA
      flexneri] align

```

Score = 42.4 bits (98), Expect = 0.029

Identities = 71/307 (23%), Positives = 124/307 (40%), Gaps = 48/307 (15%)

```

Query: 393 LNNKNSDIFDRTLPRKGLWLRVIDGH---SNQWVQGKTAPVEGYRKG VQLGG-EVFTWQN 448
      LN +   D+ D T   G W R++ G   S   +   T   G   K   +L G ++FT
Sbjct: 1010 LNKRMGDLRD-TNGEAGAWARIMSGAGSASGGYSDNYTHVQIGVDKKHELDGLDLFT--- 1065

Query: 449 ESNQLSIGLMGGQAEQRSTFHNPD TDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWM 508
      GL      +   ++      N   +G K   GAG+YA+      +GAY D
Sbjct: 1066 -----GLTMTYTD SHAS-----SNAFSGKTKSVGAGLYAS-----AIFDSGAYIDLIS 1108

Query: 509 QYQRFRRHINTED---GTERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPPAQLT 565
      +Y   +   +   GT+ ++S   +   EAGY   H T+      +++PQA+L
Sbjct: 1109 KYVHHDNEYSATFAGLGTKDYSSHSLYVGAEAGYRY---HVTEDS-----WIEPQAE LV 1159

Query: 566 YLGVNGKFSDSSENAHVNLL----GSRQLQTRVG VQAKAQFSLYKNIAIEPFAAVNALY-- 619
      Y   V+GK   D ++   +++      L R GV   FS K+   +   A   +   +
Sbjct: 1160 YGAVSGKRFDWQDRGMSVTMKDKDFNPLIGRTGVDVGKSFS-GKDWKVTARAGLG YQFDL 1218

Query: 620 --HNKPFVEMDGERRVINNKTA-IESQLGVAVKIKSHLTLQATFNRTG-KHHQAKQGA 675

```

++ + GE+R+ K I +G+ +I+ +L F + GK++
Sbjct: 1219 FANGETVLRDASGEKRIKGEKDGRILMNVLNAEIRDNLRFGLFEKSAFGKYNVDNAIN 1278
Query: 676 LNLQWTF 682
N +++F
Sbjct: 1279 ANFRYSF 1285

tr Q9L8L1 Exported serine protease SigA [sigA] [Shigella flexneri 2a] 1285 AA

align

Score = 42.4 bits (98), Expect = 0.029

Identities = 71/307 (23%), Positives = 124/307 (40%), Gaps = 48/307 (15%)

Query: 393 LNNKNSDIFDRTLPRKGLWLRVIDGH---SNQWVQKTAPVEGYRKGVLGG-EVFTWQN 448
LN + D+ D T G W R++ G S + T G K +L G ++FT
Sbjct: 1010 LNKRMGDLRD-TNGEAGAWARIMSGAGSASGGYSDNYTHVQIGVDKKHELDGLDLFT--- 1065
Query: 449 ESNQLSIGLMGGQAEQRSTFHNPDNDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWM 508
GL + ++ N +G K GAG+YA+ +GAY D
Sbjct: 1066 -----GLTMTYTDSHAS-----SNAFSGKTKSVGAGLYAS----AIFD SGAYIDLIS 1108
Query: 509 QYQRFRRHINTED---GTERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPPAQLT 565
+Y + + GT+ ++S + EAGY H T+ +++PQA+L
Sbjct: 1109 KYVHHDNEYSATFAGLGT KDYSSHSYVGAEGYRY----HVTEDS-----WIEPQAE LV 1159
Query: 566 YLGVNGKFSDSENAHVNLL---GSRQLQTRVGVQAKAQFSLYKNIAIEPFPAVNALY-- 619
Y V+GK D ++ +++ L R GV FS K+ + A + +
Sbjct: 1160 YGAVSGKRFDWQDRGMSVTMKDKDFNPLIGRTGVDVGKSFS-GKDWKVTARAGLGYQFDL 1218
Query: 620 --HNKPFVEMDGERRVINNKTA-IESQLGVAVKIKSHLTLQATFNRT-GKHHQAKQGA 675
+ + + GE+R+ K I +G+ +I+ +L F + GK++
Sbjct: 1219 FANGETVLRDASGEKRIKGEKDGRILMNVLNAEIRDNLRFGLFEKSAFGKYNVDNAIN 1278
Query: 676 LNLQWTF 682
N +++F
Sbjct: 1279 ANFRYSF 1285

tr Q7C013 Serine protease [sigA] [Shigella flexneri] 1285 AA
align

Score = 42.4 bits (98), Expect = 0.029

Identities = 71/307 (23%), Positives = 124/307 (40%), Gaps = 48/307 (15%)

Query: 393 LNNKNSDIFDRTLPRKGLWLRVIDGH---SNQWVQKTAPVEGYRKGVLGG-EVFTWQN 448
LN + D+ D T G W R++ G S + T G K +L G ++FT
Sbjct: 1010 LNKRMGDLRD-TNGEAGAWARIMSGAGSASGGYSDNYTHVQIGVDKKHELDGLDLFT--- 1065
Query: 449 ESNQLSIGLMGGQAEQRSTFHNPDNDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWM 508
GL + ++ N +G K GAG+YA+ +GAY D
Sbjct: 1066 -----GLTMTYTDSHAS-----SNAFSGKTKSVGAGLYAS----AIFD SGAYIDLIS 1108
Query: 509 QYQRFRRHINTED---GTERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPPAQLT 565


```

      +Y      +      +      GT+ ++S      +      EAGY      H T+      +++PQA+L
Sbjct: 1109 KYVHHHDNEYSATFAGLGTKDYSSHSlyVGAEAGYRY---HVTEDS-----WIEPQAE LV 1159

Query: 566 YLGVNGKFSdSENahvNLL----GSRQLQTRVGvQAkAQFSlyKNIAIEPFaAVNALY-- 619
      Y V+GK D ++ +++      L R GV      FS K+ + A + +
Sbjct: 1160 YGAVSGKRFDWQDRGMSVTMKDKDFNPLIGRTGVDVGKSFS-GKDWKVTARAGLGYQFDL 1218

Query: 620 --HNKPFgVEMdGERRVINNKTA-IESQLGVAVKIKSHLTLQATFNrQT-GKHHQAKQGA 675
      + +      + GE+R+ K I' +G+ +I+ +L      F + GK++
Sbjct: 1219 FANGETVLRDASGEKRIKGEKDGRILMNVLNAEIRDNLRFGLFEKSAFGKYNVdNAIN 1278

Query: 676 LNLQWTF 682
      N +++F
Sbjct: 1279 ANFRYSF 1285

```

```

tr Q8UJW1 Autotransporter protein [yapE] [Agrobacterium tumefaciens 868
      (strain AA
      C58 / ATCC 33970)] align

```

Score = 40.0 bits (92), Expect = 0.14

Identities = 66/293 (22%), Positives = 113/293 (38%), Gaps = 46/293 (15%)

```

Query: 411 WLRVIDGHSNqWvQgKTAP-VEGYRKGvQLGGEVFTWQNESNQLS-IGLMGGQAEQRSTF 468
      W R      ++      G +P +G G+Q G ++ + E+      GL G +
Sbjct: 599 WARTFGQNTeMKWDGTVSPSFDGNLFGLQAGQDLLGRETEAGGFDRFGLFVGYSRMNGDI 658

Query: 469 HNPDT--DNLTtGNvKGfGAGVYATWHQlQDKQtGAYADSWMQYQRfRHRINTeDGTE-R 525
      +NL G V G A W + + G Y D+ + F + G
Sbjct: 659 KGQALGWNnLAVGEVDIGGTSFGAYWTHVGAQ--GWYLDaVLMGTWfSGDATSRAGESVN 716

Query: 526 FTsKGITASIEAGYNALLAEHfTKKGNsLRvYLQPQAQLTYLGvN-----GKFS-----DS 576
      + AS+E GY L E +T      L+PQAQ+ + ++      +FS DS
Sbjct: 717 IDGSSVAASLEGGYPIALTeDWT-----LEPQAQIIWQKLSLDDeADRFSSVAFDS 767

Query: 577 ENAhvNLLGSr---QLQTRVG-VQAkAQFSlyKNIAIEPFaAVNALYhNKPFgVEMdGER 632
      +NA LG R QT G +Q + +++ + + + N P E G
Sbjct: 768 DNaVTGRLGvRLQGNyQTDsGLIqPYLKANIWHGfSSDQMTR----FDNDPIVTeTGG-- 821

Query: 633 RVINNKTaIESQLGVAVKIKSHLTLQATFNrQT---GKHHQAKQg--ALNLQW 680
      T++E G+ + +++ AT + T G+ +A +G LN++W
Sbjct: 822 -----TSLEIGGGGLVAsLTeKVSvFATVDYTTNLGGERKRAIEGNIGLNiKW 868

```

```

tr Q7D396 AGR_pAT_528p [AGR_pAT_528] [Agrobacterium tumefaciens (strain 921
      C58 / AA
      ATCC 33970)] align

```

Score = 40.0 bits (92), Expect = 0.14

Identities = 66/293 (22%), Positives = 113/293 (38%), Gaps = 46/293 (15%)

```

Query: 411 WLRVIDGHSNqWvQgKTAP-VEGYRKGvQLGGEVFTWQNESNQLS-IGLMGGQAEQRSTF 468
      W R      ++      G +P +G G+Q G ++ + E+      GL G +
Sbjct: 652 WARTFGQNTeMKWDGTVSPSFDGNLFGLQAGQDLLGRETEAGGFDRFGLFVGYSRMNGDI 711

```

Query: 469 HNPDT--DNLTGTVKGFAGVYATWHQLQDKQTGAYADSWMQYQFRHRINTEDGTE-R 525
 +NL G V G A W + + G Y D + + F + G
 Sbjct: 712 KGQALGWNNAVGEVDIGGTSFGAYWTHVGAQ--GWYLDVAVLMGTWFGSDATSRAGESVN 769

Query: 526 FTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQQAQLTYLGVN-----GKFS----DS 576
 + AS+E GY L E +T L+PQAQ+ + ++ +FS DS
 Sbjct: 770 IDGSSVAASLEGGYPIALTEDWT-----LEPQAQIIWQKLSLDDEADRFSSVAFDS 820

Query: 577 ENAHVNLLGSR---QLQTRVG-VQAKAQFSLYKNIAIEPFAAVNALYHNKPFGEVMDGER 632
 +NA LG R QT G +Q + +++ + + + N P E G
 Sbjct: 821 DNAVTGRLGVRLQGNVYQTDGLIQPYLKANIWHGFSSDQMTR----FDNDPIVTETGG-- 874

Query: 633 RVINNKTAIESQLGVAVKIKSHLTLQATFNRQT---GKHHQAKQG--ALNLQW 680
 T++E G+ + +++ AT + T G+ +A +G LN++W
 Sbjct: 875 -----TSLEIGGGLVASLTEKVSVFATVDYTTNLGGERKRAIEGNIGLNIKW 921

tr Q7WK90 Autotransporter [phg] [*Bordetella bronchiseptica* (*Alcaligenes bronchisepticus*)]

align

Score = 39.7 bits (91), Expect = 0.19

Identities = 43/213 (20%), Positives = 85/213 (39%), Gaps = 49/213 (23%)

Query: 446 WQNESNQLSIGLMGGQAEQRSTFHNPDNDNLTGTVKGFAGVYATWHQLQDKQTGAYAD 505
 W+ ++ +L +G G + R D +++ G ++ G Y T+ G Y D
 Sbjct: 204 WRLDTGRLLYGAYAGVSRARM-----DDNDIMHGRIESRFLGTYLTYVD----NGGFYVD 254

Query: 506 SWMQYQFRHRINTE-----DGTERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQ 559
 + + R ++ + D + + T S+EAGY+ L + R +++
 Sbjct: 255 AVSKLGRIDESVSFDLPLGLGDYDDDISHTTYTGVSVEAGYHFKLPQ-----RWFVE 305

Query: 560 PQAQLTYLGVNGKFSSENHAVNLLGSRQLQTRVGQAKAQFSLYKNIAIEPFAAVNALY 619
 PQAQ+ Y S ++ ++ G R GV+A F+L + P+ + + L+
 Sbjct: 306 PQAQVIY-----SRSSQTSVQG-----RAGVRAGRDFTLGGATLRPYVSASYLH 350

Query: 620 H-----NKPFGVEMDGERRVINNKTAE 642
 K + E+ G R + A++
 Sbjct: 351 EFSHDDSVDFGGKSYDAELPGSRWQLGAGAALD 383

tr Q7W8X9 Autotransporter [phg] [*Bordetella parapertussis*] 415 AA
align

Score = 39.7 bits (91), Expect = 0.19

Identities = 43/213 (20%), Positives = 85/213 (39%), Gaps = 49/213 (23%)

Query: 446 WQNESNQLSIGLMGGQAEQRSTFHNPDNDNLTGTVKGFAGVYATWHQLQDKQTGAYAD 505
 W+ ++ +L +G G + R D +++ G ++ G Y T+ G Y D
 Sbjct: 204 WRLDTGRLLYGAYAGVSRARM-----DDNDIMHGRIESRFLGTYLTYVD----NGGFYVD 254

Query: 506 SWMQYQFRHRINTE-----DGTERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQ 559
 + + R ++ + D + + T S+EAGY+ L + R +++
 Sbjct: 255 AVSKLGRIDESVSFDLPLELGDYDDDISHTTYTGVSVEAGYHFRPLPQ-----RWFVE 305

Query: 560 PQAQLTYLGVNGKFSSENHVNLLGSRQLQTRVGVQAKAQFSLYKNIAIEPFAAVNALY 619
PQAQ+ Y S ++ ++ G R GV+A F+L + P+ + + L+
Sbjct: 306 PQAQVIY-----SRSSQTSVQG-----RAGVRAGRDFTLGGATLRPYVSASYLH 350

Query: 620 H-----NKPFGEVMDGERRVINNKTAIE 642
K + E+ G R + A++
Sbjct: 351 EFSHDDSVDFGGKSYDAELPGSRWQLGAGAALD 383

tr Q84GK0 **Secreted autotransporter protein EatA [eatA] [Escherichia coli]** 1364
AA
align

Score = 39.7 bits (91), Expect = 0.19

Identities = 77/350 (22%), Positives = 136/350 (38%), Gaps = 53/350 (15%)

Query: 244 LENSLEPRWDSNVGKLKTTNADIRFNTKSE--SLLVKEDYAGGARFRFAYDPKEAKNTA 301
+ N+LW +S + +LK ++ I + L VKE A + F + +A
Sbjct: 928 MNNALWHSRNSSELKELKANDSQIELGVRGHFAKLVRKELIASNSVFLVHANNSQADQ-- 985

Query: 302 LIFEKNVTGTSDIIFENPIDDLKSLDGHQIIVNGTADKHAFRLSGKHQKGIYTLSQLQR` 361
L + G+++ I + + + +I +D++ F+ +G Q G ++ + R
Sbjct: 986 LNVTDKLQGSNNTILVDFFNKAANGTNVTITAPKGSIDENTFK-AGTQQIGFSNITPEIR 1044

Query: 362 PEGF-----LPKCKNAMIWRFMHNRLKPPIPYCALRLNNKNSDIFDRTLPRKGLW 411
E +A + + + +NN N + D
Sbjct: 1045 TENTDTATQWVLTGYQSVADARASKIATDFMDSGYKSFLTEVNNLNKRMGD----- 1095

Query: 412 LRVIDGHSNQWVQ---GKTAPVEGYRKG---VQLGGEVFTWQNESNQLSIGLMGGQAEQR 465
LR G + W + G + GYR VQ+G + ++E N I L G
Sbjct: 1096 LRDSQGDAGGARIMNGTSGESGYRDNYTHVQIGADR---KHELN--GIDLFTGALLTY 1150

Query: 466 STFHNPDNDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRF--RHRINTEDGT 523
+ N + +G K G GVYA+ ++GAY D +Y R+ +N
Sbjct: 1151 TD--NNASSQAFSGKTKSLGGGVYASGLF----ESGAYFDLIGKYLHHDNRYTLNFASLG 1204

Query: 524 ER-FTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPPQAQLTYLGVNGK 572
ER +TS + A E GY ++E+ +++PQ +L Y V+GK
Sbjct: 1205 ERSYTSLSLYAGAEIGYRYHMSSEN-----TWVEPQMELVYGSVSGK 1245

tr Q8UJX1 **Autotransporter protein [bapA] [Agrobacterium tumefaciens (strain C58 / ATCC 33970)]** 1035
AA
align

Score = 39.3 bits (90), Expect = 0.25

Identities = 49/199 (24%), Positives = 83/199 (41%), Gaps = 32/199 (16%)

Query: 409 GLWLRVIDGHSNQWVQGKTAPVEGYRKGV---QLGGEVFTWQNESNQL---SIGLMGGQ 461
G+W R+ H++ + A E Y + V Q G + + ES +L ++ G+
Sbjct: 759 GVWGRIEIVHNIHIEPRFSAAAEE-YDQNVFKLQAGIDGLLTETESGKLIGGFTVHYAHGK 817

Query: 462 AEQRSTFHNPDTNDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFRRHRINTED 521
+ RS + + + + G+G G TW+ + G Y D Q + +N+
Sbjct: 818 TDTRSVWGDGEI-----STDGYGLGGTLTWYG----ENGFYLDGQAQMTWYTSGLNSLL 867

Query: 522 GTERFTSK----GITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVN-GKFSDS 576
T G T S+E+G + ++ + PQAQL Y V+ F+D+
Sbjct: 868 ARTNLTDNNDGFGYTLSLESGTRIAIDPGWS-----VTPQAQLVYSNVDFDAFTDT 918

Query: 577 ENAHVNLLGSRQLQTRVGV 595
A V+L LQ R+GV
Sbjct: 919 FGARVSLDRGESLQGR LGV 937

tr Q7D3A2 AGR_pAT_511p [AGR_pAT_511] [Agrobacterium tumefaciens (strain 1192
C58 / AA
ATCC 33970)] align

Score = 39.3 bits (90), Expect = 0.25
Identities = 49/199 (24%), Positives = 83/199 (41%), Gaps = 32/199 (16%)

Query: 409 GLWLRVIDGHSNQVVGKTAPVEGYRKGV---QLGGEVFTWQNESNQL---SIGLMGGQ 461
G+W R+ H++ + A E Y + V Q G + + ES +L ++ G+
Sbjct: 916 GVGRIEGVHNHIEPRFSAAAAE-YDQNVFKLQAGIDGLLTETESGKLIGGFTVHYAHGK 974

Query: 462 AEQRSTFHNPDTDNLTGNGVKGFGAGVYATWHQLQDKQTGAYADSWMQYQRFRRHINTED 521
+ RS + + + + G+G G TW+ + G Y D Q + +N+
Sbjct: 975 TDTRSVWGDGEI-----STDGYGLGGTLTWYG----ENGFYLDGQAQMTWYTSGLNSLL 1024

Query: 522 GTERFTSK----GITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVN-GKFSDS 576
T G T S+E+G + ++ + PQAQL Y V+ F+D+
Sbjct: 1025 ARTNLTDNNDGFGYTLSLESGTRIAIDPGWS-----VTPQAQLVYSNVDFDAFTDT 1075

Query: 577 ENAHVNLLGSRQLQTRVGV 595
A V+L LQ R+GV
Sbjct: 1076 FGARVSLDRGESLQGR LGV 1094

tr Q6MTD5 Hypothetical transmembrane protein [MSC_0473] [Mycoplasma 873
mycoides AA
(subsp. mycoides SC)] align

Score = 38.9 bits (89), Expect = 0.32
Identities = 45/181 (24%), Positives = 74/181 (40%), Gaps = 24/181 (13%)

Query: 27 FFSILYTSPLLAVDYVYDKTKLTNDEITRLKKLRDKTSEYWKKEITYLITEDNPKVPPFPA 86
F ++ + L ++ + KTKLT ++ +LKK +DK Y K+ +IT K
Sbjct: 182 FNDLMLSFYLSRIENLNTKTKLTAKDLAKLKKFKDKYDFYLNKQEIIITNKKLK-----Q 236

Query: 87 LYPRTYQFENINNSKKISFYDQEYTEGYLVGFARGLGVAKRNGDTEEQIRKYFKECFNSN 146
L Y E +N K++ + +Y E YL + +EQI K
Sbjct: 237 LKKDQYNLELVN--KRLLKLENKY-EIYL-----NKQEIQINNNKKAKLLKK 279

Query: 147 TKIRDY-STCQAEKFGSHPLIVKSHIFSLGPKIKNSHINSEILSVGNYTEWANQVIHHIE 205
K + + C + I+ + F+ IKNS I E L ++QV+ +E
Sbjct: 280 KKSEEIKNKCLLKLENKKQKIINKNSFTKS-SIKNSLIKQEFLQARINKNTSSQVLETLE 338

Query: 206 N 206
N

Sbjct: 339 N 339

sp P45386 Immunoglobulin A1 protease precursor (EC 3.4.21.72) (IGA1 1849
IGA4_HAEIN protease) AA
[iga] [Haemophilus influenzae] align

Score = 38.5 bits (88), Expect = 0.42

Identities = 34/142 (23%), Positives = 59/142 (40%), Gaps = 20/142 (14%)

Query: 235 QTISLDEFRL-ENSLWEPRWDSNVGKLKTTNADIRFNTKSE-----SLLVKEDYAGGA 286
Q+I + L ENS W +SNV +L TN I N +++ + L +G

Sbjct: 852 QSIGTSQVNLKENSHWHLTGNSNVNQLNLTNGHIHLNAQNDANKVTTYNTLTVNSLSGNG 911

Query: 287 RFRFAYDPKEAKNTALIFEKNVTG-----TSDIIFENPIDDLKSLDGH-----QIIKV 334
F + D K+ ++ K+ TG +D E ++L D ++

Sbjct: 912 SFYYWVDFTNNKSNKVVVNKSATGNFTLQVADKTGEPNHNELTLFDASNATRNNEVTLA 971

Query: 335 NGTADKHAFRLSGKHQKGIYTL 356

NG+ D+ A++ ++ G Y L

Sbjct: 972 NGSVDRGAWKYKLRNVNGRYDL 993

tr Q9Z5R3 Phg protein (Autotransporter) [phg] [Bordetella pertussis] 418 AA

align

Score = 38.1 bits (87), Expect = 0.55

Identities = 43/213 (20%), Positives = 84/213 (39%), Gaps = 49/213 (23%)

Query: 446 WQNESNQLSIGLMGGQAEQRSTFHNPDNDLTTGNVKGFGAGVYATWHQLQDKQTGAYAD 505
W+ ++ +L +G G + R D +++ G ++ G Y T+ G Y D

Sbjct: 207 WRLDTGRLLYGAYAGVSRARM-----DDNDIMHGRIESRFLGTYLTYVD----NGGFYVD 257

Query: 506 SWMQYQFRHRINTE-----DGTERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQ 559
+ + R ++ + D + + T S EAGY+ L + R +++

Sbjct: 258 AVSKLGRIDESVSFDLPLGLGDYDDDISHTTYTGSAGYHFKLPQ-----RWFVE 308

Query: 560 PQAQLTYLGVNGKFSSENHVNLLGSRQLQTRVGVAQAQFSLYKNIAIEPFAAVNALY 619
PQAQ+ Y S ++ ++ G R GV+A F+L + P+ ++ L+

Sbjct: 309 PQAQVIY-----SRSSQTSVQG-----RAGVRAGRDFTLGGATLRPYVSASYLH 353

Query: 620 H-----NKPFGVEMDGERRVINNKTAIE 642

K + E+ G R + A++

Sbjct: 354 EFSHDDSVDFGGKSYDAELPGSRWQLGAGAALD 386

tr Q6KD18 Antigen 43 precursor Sap [sap] [Escherichia coli] 1042 AA

align

Score = 37.7 bits (86), Expect = 0.71

Identities = 38/156 (24%), Positives = 65/156 (41%), Gaps = 14/156 (8%)

Query: 417 GHSNQW--VQGKTAPVEGYRKGVLGGVFTWQNESNQLSIGLMGGQAEQRSTFHNPD TD 474
 GH N V+G T G V+L G++ + L+ G+ G A S+ D D
 Sbjct: 757 GHDNNGGIVRGATPESSGSYGVRLEGDLLRTEVAGMSLT TG VYG--AAGHSSVDVKDDD 814

Query: 475 NLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQFRHRINTEDGTERFTSKGITAS 534
 G V+ AG + L +G +AD Q R + ++++ R +G S
 Sbjct: 815 GSRAGTVRD-DAGSLGGYLN LVHTSSGLWADIVAQGRHSMKASSDNNDFRARGRGWQGS 873

Query: 535 IEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVN 570
 +E G + ++ + L+PQ Q T+ G++
 Sbjct: 874 LETGLPFSITDN-----LMLEPQLQYTWQGLS 900

tr Q8FDW4 Acreted auto transpoter toxin [sat] [Escherichia coli O6] 1299 AA

align

Score = 37.0 bits (84), Expect = 1.2

Identities = 66/290 (22%), Positives = 114/290 (38%), Gaps = 40/290 (13%)

Query: 412 LRVIDGHSNQW---VQGKTAPVEGYRKG---VQLGGVFTWQNESNQLSIGLMGGQAEQR 465
 LR I+G S W + G + G+ VQ+G + N+ + L G
 Sbjct: 1031 LRDINGESGAWARIISGTGSAGGGFSDNYTHVQVGAD-----NKHELDGLDLFTGVTMTY 1085

Query: 466 STFHNPD TDNLTTGNVKGFGAGVYATWHQLQDKQTGAYADSWMQYQFRHRINTED---G 522
 + H + +G K GAG+YA+ ++GAY D +Y + G
 Sbjct: 1086 TDSHAGS--DAFSGETKSVGAGLYAS----AMFESGAYIDLIGKYVHHDNEYTATFAGLG 1139

Query: 523 TERFTSKGITASIEAGYNALLAEHFTKKGNSLRVYLQPQAQLTYLGVNGKFSDSENAHVN 582
 T ++S A E GY H T +++PQA+L Y V+GK ++ +N
 Sbjct: 1140 TRDYSSHSWYAGAEVGYRY----HVTDSA-----WIEPQAEVYGA VSGKQFSWKDQGMN 1190

Query: 583 LL---GSRQLQTRVG VQAKAQFSLYKNIAIEPF AAVNALY-----HNKPFGVEMDGERRV 634
 L L R GV FS K+ + A + + + + + GE+R+
 Sbjct: 1191 LTMKDKDFNPLIGRTGVDVGKSFS-GKDWKVTARAGLGYQFDLFANGETVLRDASGEKRI 1249

Query: 635 INNKTA-IESQLGVAVKIKSHLT LQATFN RQT-GKHHQAKQGALNLQWTF 682
 K + +G+ +I+ +L F + GK++ N +++F
 Sbjct: 1250 KGEKDGRMLMNVGLNAEIRDNLRFGLFEKSAFGKYNVDNAINANFRYSF 1299

Database: EXPASY/UniProt

Posted date: Sep 12, 2004 1:43 PM

Number of letters in database: 501,960,298

Number of sequences in database: 1,568,464

Lambda	K	H
0.317	0.135	0.401

Gapped


Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62

Gap Penalties: Existence: 11, Extension: 1

length of query: 682
length of database: 501,960,298
effective HSP length: 131
effective length of query: 551
effective length of database: 296,491,514
effective search space: 163366824214
effective search space used: 163366824214
T: 11
A: 40
X1: 16 (7.3 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 41 (21.7 bits)
S2: 77 (34.3 bits)

 ExPASy Home page	Site Map	Search ExPASy	Contact us	Proteomics tools	Swiss-Prot
--	--------------------------	-------------------------------	----------------------------	----------------------------------	----------------------------

 ExPASy Home page	Site Map	Search ExPASy	Contact us	Swiss-Prot
--	--------------------------	-------------------------------	----------------------------	----------------------------

Search for

NiceProt View of Swiss- Prot: P19196

[Printer-friendly view](#)[Submit update](#)[Quick BlastP search](#)

[\[Entry info\]](#) [\[Name and origin\]](#) [\[References\]](#) [\[Comments\]](#) [\[Cross-references\]](#) [\[Keywords\]](#)
[\[Features\]](#) [\[Sequence\]](#) [\[Tools\]](#)

Note: most headings are clickable, even if they don't appear as links. They link to the user manual or other documents.

Entry information

Entry name	INVA_YEREN
Primary accession number	P19196
Secondary accession numbers	None
Entered in Swiss-Prot in	Release 16, November 1990
Sequence was last modified in	Release 16, November 1990
Annotations were last modified in	Release 40, October 2001

Name and origin of the protein

Protein name	Invasin
Synonyms	None
Gene name	None
From	<u>Yersinia enterocolitica</u> [TaxID: 630]
Taxonomy	<u>Bacteria</u> ; <u>Proteobacteria</u> ; <u>Gammaproteobacteria</u> ; <u>Enterobacteriales</u> ; <u>Enterobacteriaceae</u> ; <u>Yersinia</u> .

References

- [1] SEQUENCE FROM NUCLEIC ACID.
STRAIN=8081C / Serotype O:8;
MEDLINE=91041720;PubMed=2233250 [NCBI, ExPASy, EBI, Israel, Japan]
Young V.B., Miller V.L., Falkow S., Schoolnik G.K.;
"Sequence, localization and function of the invasin protein of Yersinia enterocolitica.";
Mol. Microbiol. 4:1119-1128(1990).

Comments

- **FUNCTION:** Invasin is a protein that allows enteric bacteria to penetrate cultured mammalian cells. The entry of invasin in the cell is mediated by binding several beta-1 chain integrins.
- **SUBCELLULAR LOCATION:** Outer membrane.
- **SIMILARITY:** Belongs to the intimin/invasin family.

Copyright

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a

license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)

Cross-references

EMBL X53368; CAA37448.1; -. [[EMBL](#) / [GenBank](#) / [DDBJ](#)] [[CoDingSequence](#)]
 PIR S11442; S11442.
 HSSP P11922; 1CWV. [[HSSP ENTRY](#) / [PDB](#)]
 IPR003344; Big_1.
 InterPro IPR003535; Intimin.
 IPR008964; Invasin_intimin.
[Graphical view of domain structure.](#)
 Pfam PF02369; Big_1; 1.
[Pfam graphical view of domain structure.](#)
 PRINTS PR01369; INTIMIN.
 SMART SM00634; BID_1; 1.
 ProDom [[Domain structure](#) / [List of seq. sharing at least 1 domain](#)]
 HOBACGEN [[Family](#) / [Alignment](#) / [Tree](#)]
 BLOCKS P19196.
 ProtoNet P19196.
 ProtoMap P19196.
 PRESAGE P19196.
 DIP P19196.
 ModBase P19196.
 SMR P19196; 6133FF9FDB8D9BBA.
 SWISS-2DPAGE [Get region on 2D PAGE.](#)
 UniRef View cluster of proteins with at least 50% / 90% identity.

Keywords

Outer membrane.

Features

None

Sequence information

Length: 835 Molecular weight: 91361 CRC64: 6133FF9FDB8D9BBA [This is a checksum on the
 AA Da sequence]

10	20	30	40	50	60
MYSFFNTLTV	TKIISRLILS	IGLIFGIFTY	GFSQQHYFNS	EALENPAEHN	EAFNKIISTG
70	80	90	100	110	120
TSLAVSGNAS	NITRSMVNDA	ANQEVKHWLN	RFGTTQVNVN	FDKKFSLKES	SLDWLLPWYD
130	140	150	160	170	180
SASYVFFSQL	GIRNKDSRNT	LNIGAGVRTF	QQSWMYGFNT	SYDNDMTGHN	HRIGVGAEAW
190	200	210	220	230	240
TDYLQLSANG	YFRLNGWHQS	RDFADYNERP	ASGGDIHVKA	YLPALPQLGG	KLKYEQYRGE
250	260	270	280	290	300

```

RVALFGKDNL QSNPYAVTTG LIYTPIPFIT LGVDQRMGKS RQHEIQWNLQ MDYRLGESFR
      310      320      330      340      350      360
      |      |      |      |      |      |
SQFSPAUVAG TRLLAESRYN LVERNPNIIVL EYQKQNTIKL AFSPAVL SGL PGQVYSVSAQ
      370      380      390      400      410      420
      |      |      |      |      |      |
IQSQSALQRI LWNDAQWVAA GGKLIPVSAT DYNVVLPPYK PMAPASRTVG KTGESEAAVN
      430      440      450      460      470      480
      |      |      |      |      |      |
TYTLSATAID NHGNSSNPAT LTVIVQQPQF VITSEVTDDG ALADGRTPIT VKFTVTNIDS
      490      500      510      520      530      540
      |      |      |      |      |      |
TPVAEQEGVI TTSNGALPSK VTKKTDAGQV ISIALTSFTV GVSVVTLDIQ GQQATVDVRF
      550      560      570      580      590      600
      |      |      |      |      |      |
AVLPPDVTNS SFNVSPSDIV ADGSMQSILT FVPRNKNEF VSGITDLEFI QSGVPVTISP
      610      620      630      640      650      660
      |      |      |      |      |      |
VTENADNYTA SVVGNSVGDV DITPQVGGES LDLLQKRITL YPVPKITGIN VNGEQFATDK
      670      680      690      700      710      720
      |      |      |      |      |      |
GFPKTTFNKA TFQLVMNDDV ANNTQYDWTS SYAASAPVDN QGKVNIA YKT YGSTVTVTAK
      730      740      750      760      770      780
      |      |      |      |      |      |
SKKFPSYTAT YQFKPNLWVF SGTMSLQSSV EASRNCQRTD FTALIESARA SNGSRSPDGT
      790      800      810      820      830
      |      |      |      |      |
LWGEWGSLAT YDSAEWPSGN YWTKKTSTDF VTMDMTTGDI PTSAATAYPL CAEPQ

```

P19196 in [FASTA format](#)

[View entry in original Swiss-Prot format](#)

[View entry in raw text format \(no links\)](#)

[Report form for errors/updates in this Swiss-Prot entry](#)

BLAST [BLAST submission on ExPASy/SIB](#)
or at [NCBI \(USA\)](#)



Sequence analysis tools: [ProtParam](#), [ProtScale](#),
[Compute pI/Mw](#), [PeptideMass](#), [PeptideCutter](#),
[Dotlet](#) (Java)



[ScanProsite](#), [MotifScan](#)



Search the [SWISS-MODEL Repository](#)

ExPASy Home page	Site Map	Search ExPASy	Contact us	Swiss-Prot
Hosted by NCSC US Mirror sites: Australia Bolivia Canada China Korea Switzerland Taiwan				

 ExPASy Home page	Site Map	Search ExPASy	Contact us	Swiss-Prot
--	--------------------------	-------------------------------	----------------------------	----------------------------

Search for

NiceProt

View of

Swiss-

Prot:

P11922

[Printer-friendly view](#)
[Submit update](#)
[Quick BlastP search](#)

[\[Entry info\]](#)
[\[Name and origin\]](#)
[\[References\]](#)
[\[Comments\]](#)
[\[Cross-references\]](#)
[\[Keywords\]](#)
[\[Features\]](#)
[\[Sequence\]](#)
[\[Tools\]](#)

Note: most headings are clickable, even if they don't appear as links. They link to the user manual or other documents.

Entry information

Entry name	INVA_YERPS
Primary accession number	P11922
Secondary accession numbers	None
Entered in Swiss-Prot in	Release 12, October 1989
Sequence was last modified in	Release 12, October 1989
Annotations were last modified in	Release 45, October 2004
Name and origin of the protein	
Protein name	Invasin
Synonyms	None
Gene name	None
From	<u>Yersinia pseudotuberculosis</u> [TaxID: 633]
Taxonomy	<u>Bacteria</u> ; <u>Proteobacteria</u> ; <u>Gammaproteobacteria</u> ; <u>Enterobacteriales</u> ; <u>Enterobacteriaceae</u> ; <u>Yersinia</u> .

References

- [1] SEQUENCE FROM NUCLEIC ACID.
DOI=[10.1016/0092-8674\(87\)90335-7](#);MEDLINE=87301720;PubMed=3304658 [[NCBI](#), [ExPASy](#), [EBI](#), [Israel](#), [Japan](#)]
[Isberg R.R.](#), [Voorhis D.L.](#), [Falkow S.](#);
"Identification of invasins: a protein that allows enteric bacteria to penetrate cultured mammalian cells.";
[Cell](#) 50:769-778(1987).
- [2] INTEGRIN-BINDING DOMAIN.
MEDLINE=90269235;PubMed=1693333 [[NCBI](#), [ExPASy](#), [EBI](#), [Israel](#), [Japan](#)]
[Leong J.M.](#), [Fournier R.S.](#), [Isberg R.R.](#);
"Identification of the integrin binding domain of the *Yersinia pseudotuberculosis* invasins protein.";
[EMBO J.](#) 9:1979-1989(1990).
- [3] X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 494-985.
DOI=[10.1126/science.286.5438.291](#);PubMed=10514372 [[NCBI](#), [ExPASy](#), [EBI](#), [Israel](#), [Japan](#)]
[Hamburger Z.A.](#), [Brown M.S.](#), [Isberg R.R.](#), [Bjorkman P.J.](#);

"Crystal structure of invasin: a bacterial integrin-binding protein."; Science 286:291-295(1999).

Comments

- **FUNCTION:** Invasin is a protein that allows enteric bacteria to penetrate cultured mammalian cells. The entry of invasin in the cell is mediated by binding several beta-1 chain integrins.
- **SUBCELLULAR LOCATION:** Outer surface.
- **SIMILARITY:** Belongs to the intimin/invasin family.
- **CAUTION:** It is uncertain whether Met-1, Met-17 or Met-19 is the initiator.

Copyright

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)

Cross-references

	M17448; AAA27633.1; -. [EMBL / GenBank / DDBJ] [CoDingSequence]
EMBL	M17448; AAA27632.1; ALT_INIT. [EMBL / GenBank / DDBJ] [CoDingSequence]
	M17448; AAA27634.1; ALT_INIT. [EMBL / GenBank / DDBJ] [CoDingSequence]
	M17448; AAA27635.1; ALT_INIT. [EMBL / GenBank / DDBJ] [CoDingSequence]
PIR	A29646; A29646.
PDB	1CWV; X-ray; A=494-985. [ExPASy / RCSB / EBI]
	IPR003344; Big_1.
InterPro	IPR003535; Intimin.
	IPR008964; Invasin_intimin.
	Graphical view of domain structure.
Pfam	PF02369; Big_1; 2.
	Pfam graphical view of domain structure.
PRINTS	PR01369; INTIMIN.
SMART	SM00634; BID_1; 2.
ProDom	[Domain structure / List of seq. sharing at least 1 domain]
HOBACGEN	[Family / Alignment / Tree]
BLOCKS	P11922.
ProtoNet	P11922.
ProtoMap	P11922.
PRESAGE	P11922.
DIP	P11922.
ModBase	P11922.
SMR	P11922; C875941B24BD35EE.
SWISS-2DPAGE	Get region on 2D PAGE.
UniRef	View cluster of proteins with at least 50% / 90% identity.

Keywords

3D-structure.

Features



[Feature table viewer](#)



[Feature aligner](#)

Key From To Length Description

DOMAIN	<u>494</u>	<u>985</u>	492	Extracellular.
DOMAIN	<u>494</u>	<u>594</u>	101	D1.
DOMAIN	<u>595</u>	<u>694</u>	100	D2.
DOMAIN	<u>695</u>	<u>794</u>	100	D3.
DOMAIN	<u>795</u>	<u>985</u>	191	Integrin-binding.
DOMAIN	<u>795</u>	<u>886</u>	92	D4.
DOMAIN	<u>887</u>	<u>985</u>	99	D5.
DISULFID	<u>906</u>	<u>981</u>		
STRAND	<u>503</u>	<u>509</u>	7	
STRAND	<u>513</u>	<u>514</u>	2	
STRAND	<u>521</u>	<u>528</u>	8	
TURN	<u>530</u>	<u>531</u>	2	
STRAND	<u>534</u>	<u>534</u>	1	
STRAND	<u>539</u>	<u>543</u>	5	
TURN	<u>546</u>	<u>547</u>	2	
STRAND	<u>549</u>	<u>549</u>	1	
STRAND	<u>552</u>	<u>556</u>	5	
TURN	<u>559</u>	<u>560</u>	2	
STRAND	<u>562</u>	<u>568</u>	7	
STRAND	<u>573</u>	<u>581</u>	9	
TURN	<u>582</u>	<u>583</u>	2	
STRAND	<u>584</u>	<u>593</u>	10	
STRAND	<u>597</u>	<u>597</u>	1	
HELIX	<u>599</u>	<u>601</u>	3	
STRAND	<u>603</u>	<u>607</u>	5	
STRAND	<u>610</u>	<u>612</u>	3	
TURN	<u>613</u>	<u>614</u>	2	
STRAND	<u>619</u>	<u>624</u>	6	
STRAND	<u>626</u>	<u>627</u>	2	
TURN	<u>628</u>	<u>629</u>	2	
STRAND	<u>632</u>	<u>632</u>	1	
STRAND	<u>638</u>	<u>642</u>	5	
STRAND	<u>646</u>	<u>647</u>	2	
STRAND	<u>651</u>	<u>652</u>	2	
STRAND	<u>657</u>	<u>663</u>	7	
STRAND	<u>668</u>	<u>676</u>	9	
TURN	<u>677</u>	<u>678</u>	2	
STRAND	<u>679</u>	<u>691</u>	13	
TURN	<u>698</u>	<u>700</u>	3	
STRAND	<u>702</u>	<u>705</u>	4	
STRAND	<u>709</u>	<u>711</u>	3	
TURN	<u>712</u>	<u>713</u>	2	
STRAND	<u>718</u>	<u>723</u>	6	
STRAND	<u>725</u>	<u>725</u>	1	
TURN	<u>727</u>	<u>728</u>	2	
STRAND	<u>731</u>	<u>731</u>	1	
STRAND	<u>738</u>	<u>744</u>	7	
STRAND	<u>748</u>	<u>749</u>	2	

STRAND	<u>753</u>	<u>754</u>	2
STRAND	<u>759</u>	<u>764</u>	6
STRAND	<u>769</u>	<u>777</u>	9
TURN	<u>778</u>	<u>779</u>	2
STRAND	<u>780</u>	<u>781</u>	2
HELIX	<u>783</u>	<u>785</u>	3
STRAND	<u>787</u>	<u>792</u>	6
STRAND	<u>797</u>	<u>802</u>	6
TURN	<u>803</u>	<u>804</u>	2
STRAND	<u>805</u>	<u>807</u>	3
TURN	<u>809</u>	<u>810</u>	2
STRAND	<u>816</u>	<u>818</u>	3
TURN	<u>819</u>	<u>820</u>	2
STRAND	<u>822</u>	<u>827</u>	6
TURN	<u>828</u>	<u>829</u>	2
HELIX	<u>831</u>	<u>836</u>	6
STRAND	<u>837</u>	<u>841</u>	5
TURN	<u>844</u>	<u>846</u>	3
STRAND	<u>847</u>	<u>849</u>	3
TURN	<u>851</u>	<u>852</u>	2
STRAND	<u>854</u>	<u>857</u>	4
STRAND	<u>865</u>	<u>871</u>	7
TURN	<u>875</u>	<u>876</u>	2
STRAND	<u>878</u>	<u>883</u>	6
STRAND	<u>887</u>	<u>891</u>	5
STRAND	<u>897</u>	<u>897</u>	1
HELIX	<u>899</u>	<u>905</u>	7
TURN	<u>906</u>	<u>906</u>	1
TURN	<u>910</u>	<u>911</u>	2
STRAND	<u>912</u>	<u>914</u>	3
STRAND	<u>916</u>	<u>916</u>	1
HELIX	<u>917</u>	<u>920</u>	4
HELIX	<u>931</u>	<u>935</u>	5
HELIX	<u>938</u>	<u>940</u>	3
TURN	<u>941</u>	<u>941</u>	1
TURN	<u>943</u>	<u>944</u>	2
STRAND	<u>949</u>	<u>953</u>	5
STRAND	<u>959</u>	<u>963</u>	5
TURN	<u>964</u>	<u>966</u>	3
STRAND	<u>969</u>	<u>970</u>	2
STRAND	<u>977</u>	<u>985</u>	9

Sequence information

Length: **985** Molecular weight: **106627** CRC64: **C875941B24BD35EE** [This is a checksum on the
AA **Da** sequence]

10	20	30	40	50	60
MVFQPISEFL	LIRNAGMSMY	FNKIISFNII	SRVICIFLI	CGMFMAGASE	KYDANAPQQV

70	80	90	100	110	120
QPYSVSSSAF	ENLHPNNE	SSINPFSASD	TERNAAIIDR	ANKEQETEAV	NKMISTGARL
130	140	150	160	170	180
AASGRASDVA	HSMVGDAVNQ	EIKQWLNRF	TAQVNLNFDK	NFSLKESSLD	WLAPWYDSAS
190	200	210	220	230	240
FLFFSQLGIR	NKDSRNTLNL	GVGIRTLENG	WLYGLNTFYD	NDLTGHNHRI	GLGAEAWTDY
250	260	270	280	290	300
LQLAANGYFR	LNGWHSSRDF	SDYKERPATG	GDLRANAYLP	ALPQLGGKLM	YEQYTGERVA
310	320	330	340	350	360
LFGKDNLQRN	PYAVTAGINY	TPVPLLTGV	DQRMGKSSKH	ETQWNLMQNY	RLGESFQSQL
370	380	390	400	410	420
SPSAVAGTRL	LAESRYNLVD	RNNNIVLEYQ	KQQVVKLTLS	PATISGLPGQ	VYQVNAQVQG
430	440	450	460	470	480
ASAVREIVWS	DAELIAAGGT	LTPLSTTQFN	LVLPPYKRTA	QVSRVTDDL	ANFYSLSALA
490	500	510	520	530	540
VDHQGNRSNS	FTLSVTVQQP	QLTLTAAVIG	DGAPANGKTA	ITVEFTVADF	EGKPLAQEV
550	560	570	580	590	600
VITTNNGALP	NKITEKTDAN	GVARIALTNT	TDGVTVVTAE	VEGQRQSVDT	HFVKGTIAAD
610	620	630	640	650	660
KSTLAAVPTS	IIADGLMAST	ITLELKDTYG	DPQAGANVAF	DTTLGNMGVI	TDHNDGTYS
670	680	690	700	710	720
PLTSTTLGVA	TVT VKVDGAA	FSVPSVTVNF	TADPIPDAGR	SSFTVSTPDI	LADGTMSSTL
730	740	750	760	770	780
SFVPVDKNGH	FISGMQGLSF	TQNGVPVSIS	PITEQPDSYT	ATVVGNSVGD	VTITPQVDTL
790	800	810	820	830	840
ILSTLQKKIS	LFPVPTLTGI	LVNGQNFATD	KGFPKTIFKN	ATFQLQMDND	VANNTQYEWS
850	860	870	880	890	900
SSFTPNVSVN	DQGQVTITYQ	TYSEVAVTAK	SKKFPSYSVS	YRFYPNRWIY	DGGRSLVSSL
910	920	930	940	950	960
EASRQCQGS	MSAVLESSRA	TNGTRAPDGT	LWGEWGSLLTA	YSSDWQSGEY	WVKKTSTDFF

970 980
| |
TMNMDTGALQ PGPAYLAFPL CALSI

P11922 in [FASTA](#)
[format](#)

[View entry in original Swiss-Prot format](#)

[View entry in raw text format \(no links\)](#)

[Report form for errors/updates in this Swiss-Prot entry](#)

BLAST [BLAST submission on](#)
[ExPASy/SIB](#)
or at [NCBI \(USA\)](#)




Sequence analysis tools: [ProtParam](#), [ProtScale](#),
[Compute pI/Mw](#), [PeptideMass](#), [PeptideCutter](#),
[Dotlet \(Java\)](#)



[ScanProsite](#), [MotifScan](#)



Search the [SWISS-MODEL Repository](#)

 ExPASy Home page	Site Map	Search ExPASy	Contact us	Swiss-Prot				
Hosted by NCSC US	Mirror sites:	Australia	Bolivia	Canada	China	Korea	Switzerland	Taiwan

WEST Search History

DATE: Tuesday, September 14, 2004

Hide?	<u>Set</u> <u>Name</u>	<u>Query</u>	<u>Hit</u> <u>Count</u>
		<i>DB=PGPB,USPT,USOC,EPAB,JPAB,DWPI,TDBD; PLUR=YES; OP=AND</i>	
<input type="checkbox"/>	L1	ipa\$1	29122
<input type="checkbox"/>	L2	L1 same (coli or shigella or salmonella or yersinia or rickettsia or brucella or erhlichiae or edwarsiella or campylobacter or legionella or neisseria)	479
<input type="checkbox"/>	L3	L2 and (invasin or intimin or yop)	71
<input type="checkbox"/>	L4	L3 and (lps or lipo-polysaccharide or lipopolysaccharide)	34
<input type="checkbox"/>	L5	('6245892' '20020197276' '6083683' '6680374' '6277379' '20010009957')!.PN.	10

END OF SEARCH HISTORY

Search Results - Record(s) 1 through 10 of 10 returned.

L5: Entry 1 of 10	File: PGPB	Dec 26, 2002
PGPUB-DOCUMENT-NUMBER: 20020197276 DOCUMENT-IDENTIFIER: US 20020197276 A1		
TITLE: Heterologous protection induced by immunization with invaplex vaccine PUBLICATION-DATE: December 26, 2002		
US-CL-CURRENT: <u>424/203.1</u> INT-CL: [07] <u>A61 K 39/116</u>		

L5: Entry 2 of 10	File: PGPB	Jul 26, 2001
PGPUB-DOCUMENT-NUMBER: 20010009957 DOCUMENT-IDENTIFIER: US 20010009957 A1		
TITLE: Invaplex from gram negative bacteria, method of purification and methods of use PUBLICATION-DATE: July 26, 2001		
US-CL-CURRENT: <u>530/395</u> ; <u>435/195</u> INT-CL: [07] <u>C12 N 9/14</u> , <u>C07 K 14/24</u>		

L5: Entry 3 of 10	File: USPT	Jan 20, 2004
US-PAT-NO: 6680374 DOCUMENT-IDENTIFIER: US 6680374 B2		
TITLE: Invaplex from gram negative bacteria, method of purification and methods of use		
DATE-ISSUED: January 20, 2004		
US-CL-CURRENT: <u>530/388.1</u> ; <u>424/130.1</u> , <u>424/141.1</u> , <u>424/150.1</u> , <u>424/164.1</u> , <u>435/329</u> , <u>435/332</u> , <u>435/340</u> , <u>530/350</u> , <u>530/388.2</u> , <u>530/388.4</u>		
INT-CL: [07] <u>C07 K 16/00</u> , <u>C12 P 21/08</u>		

L5: Entry 4 of 10	File: USPT	Aug 21, 2001
US-PAT-NO: 6277379 DOCUMENT-IDENTIFIER: US 6277379 B1		
TITLE: Use of purified invaplex from gram negative bacteria as a vaccine		
DATE-ISSUED: August 21, 2001		
US-CL-CURRENT: <u>424/197.11</u> ; <u>424/193.1</u> , <u>424/203.1</u> , <u>424/234.1</u> , <u>424/241.1</u> , <u>424/249.1</u> , <u>424/252.1</u> , <u>424/258.1</u> , <u>435/975</u> , <u>530/350</u> , <u>536/123.1</u>		

INT-CL: [07] A61 K 39/385

L5: Entry 5 of 10

File: USPT

Jun 12, 2001

US-PAT-NO: 6245892

DOCUMENT-IDENTIFIER: US 6245892 B1

TITLE: Invaplex from gram negative bacteria, method of purification and methods of use

DATE-ISSUED: June 12, 2001

US-CL-CURRENT: 530/350; 424/282.1, 435/7.2, 530/416

INT-CL: [07] C07 K 14/00, C07 K 1/00, G01 N 33/53, A61 K 45/00

L5: Entry 6 of 10

File: USPT

Jul 4, 2000

US-PAT-NO: 6083683

DOCUMENT-IDENTIFIER: US 6083683 A

TITLE: Methods for detecting shigella bacteria or antibodies to shigella bacteria with an immunoassay

DATE-ISSUED: July 4, 2000

US-CL-CURRENT: 435/4; 424/282.1, 424/93.4, 435/252.1, 435/29, 435/34, 435/822, 435/975

INT-CL: [07] C12 Q 1/00

L5: Entry 7 of 10

File: DWPI

Dec 3, 2002

DERWENT-ACC-NO: 2003-129359

ABSTRACTED-PUB-NO: WO 200294190A

COPYRIGHT 2004 DERWENT INFORMATION LTD

TITLE: Inducing in a subject a protective immune response against infection with a first invasive gram-negative bacteria by administering a composition comprising Invaplex 50 from a second heterologous invasive gram negative bacteria

INT-CL (IPC): A61 K 0/00, A61 K 39/02, A61 K 39/108, A61 K 39/112, A61 K 39/116, A61 K 39/385, C07 K 1/00, G01 N 33/53

Derwent-CL (DC): B04, D16

CPI Codes: B04-B04C1; B04-C02V; B04-F10A; B04-F10A3; B04-F10A5; B04-N03; B14-A01A; B14-A01A3; B14-A01A5; B14-S11B; D05-H07;

L5: Entry 8 of 10

File: DWPI

Aug 25, 2004

DERWENT-ACC-NO: 2000-292989

ABSTRACTED-PUB-NO: US 6277379B

COPYRIGHT 2004 DERWENT INFORMATION LTD

TITLE: New vaccine for protection against infection with gram-negative bacteria comprising Invaplex of the bacteria to elicit protective antibodies

INT-CL (IPC): A61 K 0/00, A61 K 39/02, A61 K 39/385, A61 P 31/04

Derwent-CL (DC): B04, D16

CPI Codes: B04-B04M; B04-F10A; B04-G07; B11-C06; B14-A01A; B14-S11B; D05-H07; D05-H11;

L5: Entry 9 of 10

File: DWPI

Jan 20, 2004

DERWENT-ACC-NO: 2000-292988

ABSTRACTED-PUB-NO: US 6245892B

COPYRIGHT 2004 DERWENT INFORMATION LTD

TITLE: New composition comprising isolated Invaplex of gram-negative bacteria comprising at least one invasin protein associated with LPS of the gram-negative bacteria

INT-CL (IPC): A61 K 0/00, A61 K 45/00, C07 K 1/00, C07 K 14/00, C07 K 14/24, C07 K 16/00, C12 N 9/14, C12 P 21/08, G01 N 33/53

Derwent-CL (DC): B04, D16

CPI Codes: B04-B04L; B04-B04M; B04-F01; B04-F10A; B04-N03; D05-H04; D05-H07; D05-H09; D05-H11;

L5: Entry 10 of 10

File: DWPI

Apr 7, 2003

DERWENT-ACC-NO: 1996-209840

ABSTRACTED-PUB-NO: US 5679564A

COPYRIGHT 2004 DERWENT INFORMATION LTD

TITLE: Prodn. of enhanced antigenic enteric bacteria for use in e.g. vaccines - uses bile acids, bile salts, oxygen, nitrogen, carbon-di:oxide and divalent cation chelant 704283 B

INT-CL (IPC): A01 N 63/00, A61 K 39/00, A61 K 39/02, A61 K 39/106, A61 K 39/108, A61 K 39/38, A61 K 45/00, A61 P 1/00, A61 P 31/04, C12 N 0/00, C12 N 1/00, C12 N 1/12, C12 N 1/20, C12 N 1/38, C12 Q 1/00, C12 Q 1/02, C12 Q 1/04, G01 N 33/53, G01 N 33/531, G01 N 33/569, C12 N 1/20, C12 R 1:19, C12 N 1/20, C12 R 1:63, C12 N 1/20, C12 R 1:42, C12 N 1/20, C12 R 1:01, C12 N 1/20, C12 R 1:19, C12 N 1/20, C12 R 1:63, C12 N 1/20, C12 R 1:42, C12 N 1/20, C12 R 1:01

Derwent-CL (DC): B04, C07, D16, S03

CPI Codes: B04-F10A; C04-F10A; B04-G01; C04-G01; B10-B01B; C10-B01B; B11-C07A; C11-C07A; B12-K04A; C12-K04A; B14-S11B; C14-S11B; D05-H04; D05-H07; D05-H09;

EPI Codes: S03-E14H4;

[Previous Page](#)

[Next Page](#)



Journal of Bacteriology

[▶ HOME](#)

T. Taniya, J. Mitobe, S.-i. Nakayama, Q. Mingshan, K. Okuda, and H. Watanabe

Determination of the InvE Binding Site Required for Expression of IpaB of the *Shigella sonnei* Virulence Plasmid: Involvement of a ParB BoxA-Like Sequence

J. Bacteriol., September 1, 2003; 185(17): 5158 - 5165.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)



Infection and Immunity

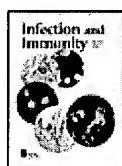
[▶ HOME](#)

A. T. Harrington, P. D. Hearn, W. L. Picking, J. R. Barker, A. Wessel, and W. D. Picking

Structural Characterization of the N Terminus of IpaC from *Shigella flexneri*

Infect. Immun., March 1, 2003; 71(3): 1255 - 1264.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)



Infection and Immunity

[▶ HOME](#)

B. Foutier, P. Troisfontaines, D. Vertommen, M.-N. Marenne, M. Rider, C. Parsot, and G. R. Cornelis

Identification of Substrates and Chaperone from the *Yersinia enterocolitica* 1B Ysa Type III Secretion System

Infect. Immun., January 1, 2003; 71(1): 242 - 253.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)



Applied and Environmental Microbiology

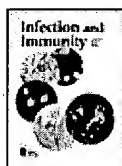
[▶ HOME](#)

S. M. Faruque, R. Khan, M. Kamruzzaman, S. Yamasaki, Q. S. Ahmad, T. Azim, G. B. Nair, Y. Takeda, and D. A. Sack

Isolation of *Shigella dysenteriae* Type 1 and *S. flexneri* Strains from Surface Waters in Bangladesh: Comparative Molecular Analysis of Environmental *Shigella* Isolates versus Clinical Strains

Appl. Envir. Microbiol., August 1, 2002; 68(8): 3908 - 3913.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)



Infection and Immunity

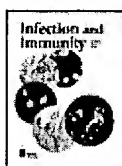
[▶ HOME](#)

M. M. Venkatesan, M. B. Goldberg, D. J. Rose, E. J. Grotbeck, V. Burland, and F. R. Blattner

Complete DNA Sequence and Analysis of the Large Virulence Plasmid of *Shigella flexneri*

Infect. Immun., May 1, 2001; 69(5): 3271 - 3285.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)



Infection and Immunity

[▶ HOME](#)

C. M. Fernandez-Prada, D. L. Hoover, B. D. Tall, A. B. Hartman, J. Kopelowitz, and M. M. Venkatesan

***Shigella flexneri* IpaH7.8 Facilitates Escape of Virulent Bacteria from the Endocytic Vacuoles of Mouse and Human Macrophages**

Infect. Immun., June 1, 2000; 68(6): 3608 - 3619.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)



Gut [▶ HOME](#)

P J Jenks, F Megraud, and A Labigne

Clinical outcome after infection with *Helicobacter pylori* does not appear to be reliably predicted by the presence of any of the genes of the cag pathogenicity island

Gut, December 1, 1998; 43(6): 752 - 758.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)



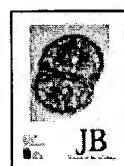
Antimicrobial Agents and Chemotherapy [▶ HOME](#)

J. L. Kadurugamuwa and T. J. Beveridge

Delivery of the Non-Membrane-Permeative Antibiotic Gentamicin into Mammalian Cells by Using *Shigella flexneri* Membrane Vesicles

Antimicrob. Agents Chemother., June 1, 1998; 42(6): 1476 - 1483.

[\[Abstract\]](#) [\[Full Text\]](#)



Journal of Bacteriology [▶ HOME](#)

S.-i. Nakayama and H. Watanabe

Identification of cpxR as a Positive Regulator Essential for Expression of the *Shigella sonnei* virF Gene

J. Bacteriol., July 15, 1998; 180(14): 3522 - 3528.

[\[Abstract\]](#) [\[Full Text\]](#)



Microbiology and Molecular Biology Reviews [▶ HOME](#)

C. J. Hueck

Type III Protein Secretion Systems in Bacterial Pathogens of Animals and Plants

Microbiol. Mol. Biol. Rev., June 1, 1998; 62(2): 379 - 433.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)



Infection and Immunity [▶ HOME](#)

K. R. Turbyfill, J. A. Mertz, C. P. Mallett, and E. V. Oaks

Identification of Epitope and Surface-Exposed Domains of *Shigella flexneri* Invasion Plasmid Antigen D (IpaD)

Infect. Immun., May 1, 1998; 66(5): 1999 - 2006.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)



Infection and Immunity [▶ HOME](#)

G. J. Leclerc, C. Tartera, and E. S. Metcalf

Environmental Regulation of *Salmonella typhi* Invasion-Defective Mutants

Infect. Immun., February 1, 1998; 66(2): 682 - 691.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

[Current Issue](#) | [Archives](#) | [Online Submission](#) | [Info for Authors](#) | [Subscribe](#) | [About](#) | [Editorial Board](#) | [Contact](#) | [Sitemap](#)

[Copyright © 1988 by the National Academy of Sciences](#)